CONTROL OF MOVEMENT IN DECISION MAKING: MOVEMENT VIGOR, REACTION, AND THEIR RELATIONSHIP WITH SUBJECTIVE VALUE AND EFFORT

by

Tehrim Yoon

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

October 2019

Abstract

Decision-making and motor control both involve dopaminergic neurons in the basal ganglia. Thus, it seems logical to think that these two aspects are strongly related. There have been many studies showing that movements are modulated by reward, and even by dopaminergic activity. Recently, some studies have shown that even saccades, which were considered to have an internal value of baseline velocity for each subject, are modulated by reward.

In Chapter 2, we first explain how we can modulate effort with eccentricity. We show that eye fixation on a location requires motor effort depending on its eccentricity. We also show that the velocity of saccade made from target to target is not decided by the difference in reward between the two targets, but rather by the sum of rewards.

In Chapter 3, we construct a model to explain both gaze duration and saccade vigor by exploiting an existing model called the marginal value theorem (MVT), which originally only explained how harvest duration changes. We generated a task environment for eye movements that is analogous to the foraging environment and were able to predict the behavioral pattern of gaze duration from our new model. In addition to local factors, we also observed the effects of past and future parameters on gaze duration, which is only predicted when the global utility is defined as the sum of all gains divided by the sum of all times. We also observed modulation of saccade vigor by a similar set of parameters.

In Chapter 4, we try to explain how saccade vigor and reaction time to a cue are related to the utility value associated with the target. We constructed a task that was a mixture of cued response trials and decision-making trials. We observed that not only were vigor and reaction

time modulated by reward, but also that they were more strongly modulated by utility values derived from the decision-making trials. Saccade vigor and reaction time were also shown to be dependent on the signed utility value, not on the salience (unsigned) of the associated stimulus.

In conclusion, we show that motor control of saccades is closely related to decision making.

Primary Reader and Advisor: Reza Shadmehr

Secondary Reader: Jeremiah Y. Cohen

Thesis Committee: Marshall H. Shuler, Jeremiah Y. Cohen, and David S. Zee

Acknowledgements

I greatly thank my advisor, Professor Reza Shadmehr, who introduced me to the field of computational motor control. I also thank him for his guidance along the road to my Ph.D. degree. Reza's exceptional intuition and experience in this wonderful field helped me immensely to pursue research in this field, and together we have found considerable success. He also guided me to overcome a number of difficulties and failures while still providing me with valuable opportunities to improve my knowledge, skills, and creativity. His expertise and guidance also provided me with a lot of experience in various areas, which helped me to widen and consolidate my research interests.

I am also grateful to Professor Jeremiah Cohen for always being supportive yet critical of my research. His experience and expertise in the neural process of decision making helped a great deal in setting up our hypotheses. I also thank Professor Marshall Shuler, another member of my thesis advisory committee, who aided me with his expertise in the cognitive process of decision making, and Professor David Zee, also a member of my committee, who helped me greatly with the process of saccades. I also thank Professor Jim Knierim, Professor Amy Bastian, Professor Dwight Bergles, Professor King-Wai Yau, and the rest of the steering committee for helping me out in the most difficult initial years of the program.

I would also like to thank all the members of the Shadmehr lab. Everyone has been friendly and helped immensely with my research and degree program.

Finally, I dedicate this to my family for their everlasting love and support, notably during my several years in Baltimore.

Contents

List of Figures

1. Introduction

1.1. Saccades

Saccades are rapid movements of both eyes from one fixation to another. Humans make saccades in situations such as reading or looking at pictures. This rapid eye movement is generally known to happen in order to change input of the visual map on the retina, which has differing concentrations of photoreceptors and thus provides an advantage when humans want to focus on a particular area of the visual map. We are interested in voluntary saccades, which are saccades made voluntarily to shift the visual map in order to improve resolution in an area of interest and acquire more information.

Saccades have a stereotypical bell-shaped velocity profile. They are usually asymmetric: they accelerate faster to peak velocity than they decelerate after the peak, therefore skewing the shape to the right (positive skewness). As a result, the position profile of a saccade is S-shaped. Voluntary saccades have a typical reaction time >200ms, and saccades that have a shorter reaction time are considered reflexive and are called express saccades (Fischer and Boch, 1983). Saccade amplitude is usually shown in units of degrees, indicating how much the eye rotated around its axes.

A saccade is generally faster when its amplitude is larger, but the steepness rapidly declines after 20°, and its peak velocity usually asymptotes below 1000°/s (Bahill et al., 1975). In fact, when tested for an abundant number of saccades of varying amplitude, each person shows a systematic relationship between amplitude and peak velocity, and the initial discoverers designated this the “main sequence” of saccades (Bahill et al., 1975). The authors also reported

that variability in saccade vigor is mostly accounted for by amplitude and not by its starting position or direction. In fact, most people believed that saccade vigor is generally invariant with respect to external factors (excluding, of course, factors such as fatigue) since humans cannot consciously control saccade vigor. We will investigate this more closely and describe various recent studies that refute this concept in Section 1.5.

1.2. Neural basis of saccadic behavior

Although saccade data is easier to analyze compared to the likes of limb movements due to smaller number of dimension and its stereotypical velocity profile, they still involve a number of brain areas and the functionality of these areas has still not been completely explained. Saccades in various directions involve slightly different circuitry, partly because they use different muscles and thus different cranial nerves. In this thesis we will only consider horizontal saccades.

The initial motor command comes from burst generators located within the brainstem. These neurons are called excitatory burst neurons (EBN) and inhibitory burst neurons (IBN), respectively, in areas called the paramedian pontine reticular formation (PPRF) for EBNs and medullary reticular formation (MRF) for IBNs (Strassman et al., 1986a, 1986b). These neurons are essentially silent during fixation but show bursting activity during saccades, with a short lead time of several milliseconds (Van Gisbergen et al., 1981). Each burst generator fires differently for each direction, and the most active direction is called the on-direction (as opposed to off-direction) and saccades in such a direction are called ipsiversive (as opposed to contraversive). EBNs fire to excite motor neurons that cause the eyes to move in the ipsiversive direction, while

IBNs fire to inhibit motor neurons to prevent moving in the contraversive direction. This coordination between two burst neurons results in a properly conjugated saccade.

Holding the eyes still was first discovered to be controlled by omnipause neurons (OPN), which are located in the nucleus raphe interpositus (part of the PPRF) in the brainstem and are known to inhibit both types of burst neurons and thus prevent eye movement (Keller, 1974). However, holding the eye in a certain position requires the existence of another signal that encodes the position. In fact, David Robinson proposed a model for the control of saccades that included a "neural integrator" component, which he hypothesized would integrate movement signals over time and add it to the movement signal (Robinson, 1973). The integrator was soon discovered by Robinson himself in a region called the nucleus prepositus hypoglossi (NPH, or simply the prepositus) by observing that injecting neurotoxins into this region did not disrupt saccades but caused the eyes to rotate back to near center (Cannon and Robinson, 1987). The output from the prepositus also goes to the motor neurons, which project to oculomotor neurons (cranial nerve III) and thus onto the muscles, generating tension to hold the eye in the position.

The burst neurons get input from the superior colliculus (SC), which receives signals from eye-related areas of the cortex. The dorsal layer receives input from other areas and forms a visual map, while the ventral layer is organized by intended saccade amplitude and generates appropriate motor signals. A number of brain regions also elicit eye movements, including the frontal eye field (FEF) (Robinson and Fuchs, 1969) and supplemental eye field (SEF) (Schlag and Schlag-Rey, 1987). This connection also involves the basal ganglia (Parthasarathy et al., 1992). The implications of involvement of basal ganglia will be discussed in Section 1.4.

1.3. Decision and utility

Decision making is a cognitive process of selection of actions among several alternative options. In fact, microeconomics is the field that specifically studies the behavior of making decisions. Economists use a concept called a utility function, which is an amount of satisfaction, worth, or value of an alternative. This concept is based on the assumption that it is possible to assign real numbers to describe preference, so that an individual would prefer the alternative that has the highest utility to the others (von Neumann & Morgenstern, 1944). The value is subjective (so each individual has a different utility for the same alternative) and is a nonlinear function by amount of gain and thus needs to be estimated each time (Kahneman & Tversky, 1979). However, it is generally higher for larger reward and smaller cost or effort.

Traditionally, utility is estimated using each individual's behavior during decision making. The most direct way would be to simply ask each subject to manually assign a value to each alternative (cardinal utility) (Grether and Plott, 1979). But this is not practical with nonhuman animals, and sometimes the reported numbers from human subjects do not explain their actual choice behavior. In order to overcome this limitation, researchers generally try to measure the certainty equivalent (CE). The CE refers to a single value that a subject would choose over a risky, unknown target option (thus indicating that they have same utility). Measuring CE for a number of various options will eventually reveal the utility function.

Theoretically, estimate of CE could be found by measuring probabilities of choice, which is measured by providing a subject with same decision trial repetitively. For a target option whose CE is unknown, probabilities should be measured over a number of different single value against the target option in a two-alternative choice task. A logistic fit over the probabilities will give CE, but this method requires a lot of trials just to measure CE of one target option. To get the most accurate estimates of CE, adaptive methods (trials change based on an individual's

choice) are used, such as PEST (Pollack, 1968). PEST is a way of varying the competing single value against the target option based on the subject's choice, rather than assigning it before the experiment (thus the term "adaptive"). This method requires a smaller number of trials to measure each CE and is currently widely used. However, transitioning CE for multiple options to estimate the utility curve is not an easy task. In Chapter 4, we suggest a new way to measure the utility curve in a more efficient way.

1.4. Neural basis of reward and its relationship to the neural basis of saccadic behavior

Ever since it was discovered that dopamine encodes reward prediction error (Schultz et al., 1997), the dopaminergic system in the midbrain has been the focus as the main region of interest for decision making. Two regions were mainly acknowledged; the ventral tegmental area (VTA) of the midbrain (mesolimbic pathway) and the substantia nigra pars compacta (SNc) (nigrostriatal pathway) (Ungelstedt, 1971). Both areas innervate the striatum, which is a crucial part of the basal ganglia. Both VTA and SNc dopaminergic neurons are associated with reward prediction error (Schultz et al., 1997); in fact, these two areas are closely located, and their roles are not dichotomous, according to recent discoveries (Watabe-Uchida et al., 2012; Ilango et al., 2014).

The VTA is the main region of interest in decision making studies due to its relationship with the frontal area of the cortex. The VTA projects to the nucleus accumbens (NAcc) which forms the ventral striatum, and also innervates the ventral pallidum, which is part of the basal ganglia (Beckstead et al., 1979; Heimer, 1978). The ventral pallidum projects to the medial

dorsal (MD) thalamus, which also receives inputs from the prefrontal cortical areas such as the orbitofrontal cortex (OFC), thus forming a pathway (Young et al., 1984). The prefrontal areas are known to encode various forms of reward and cost (Matsumoto et al., 2003; Amiez et al., 2005; Roesch et al., 2006), and thus have been thoroughly investigated for its crucial role in decision making. In fact, these cortical regions also innervate the striatum, thus playing an important role in movement.

On the other hand, the SNc is directly associated with the basal ganglial pathway that is involved with controlling movements. The striatal neurons have dopamine receptors that receive signals from the SNc, thus regulating the input from the motor cortices (Gerfen et al., 1987) and playing an important role on behavior selection (Cools, 1980). Two different receptors in the striatal neurons, D1 and D2, cause separate effects (Gerfen et al., 1990). The D1 receptor neurons project to the substantia nigra pars reticulata (SNr), which projects to the thalamus and the superior colliculus (direct pathway). Because these two neurons are both inhibitory, this pathway results in an excitatory output. Meanwhile, the D2 receptor neurons project to the globus pallidum externus (GPe), then to the globus pallidum internus (GPi) (which projects to the thalamus) through an intermediate region called the subthalamic nucleus (STN) (indirect pathway). This contains three inhibitory connections and thus result in an inhibitory output. There is some evidence that the direct pathway may be involved in learning from rewards, while the indirect pathway may be involved in learning from punishments (Hikida et al., 2010; Kravitz et al., 2012). Based on the vast interconnection between the basal ganglia and dopaminergic neurons, it seems obvious that these two regions are closely related.

We focus on two known properties of the dopaminergic system and the potential relationship between motor control and reward circuit. First, loss of dopamine usually results in a

degenerative disorder called Parkinson's disease, which results in motor disorders such as rigidity, shakiness, and slow movement. Parkinson's disease patients are also known to display some impairment in decision making, such as reduced reward sensitivity in gambling tasks (Mimura et al., 2006) or speed-accuracy tradeoff (Manohar et al., 2015). Second, while basal ganglia are involved in the control of movements, these neurons are known to encode reward and cost in a decision-making task, although in a limited way (Pasquereau and Turner, 2013). As a result, these properties suggest that utility, which is decided by the subjective evaluation of reward and effort and also determines decision-making activity, could also modulate control of movement.

In fact, several studies show that basal ganglia are involved in the control of saccades. As mentioned earlier, the caudate nucleus (CN) is part of the basal ganglia that receives input from the aforementioned dopaminergic regions and sends output to the substantia nigra pars reticulata (SNr) and the globus pallidum externus (GPe). The firing rate of CN neurons has been shown to be related to speed and reaction time modulation by expected reward (Kawagoe et al., 1998) and also this relationship existed when expected reward was controlled for some neurons (Itoh et al., 2003). For the control of saccades, the SNr sends inhibitory output to the SC (Hikosaka & Wurtz, 1983) and shows bursting and pausing activity that generates saccades (Handel and Glimcher, 1999). Another pathway includes the VTA of the thalamus and the ventral pallidum which has also been shown to encode expected reward and modulate motion (Tachibana and Hikosaka, 2012). In summary, neural activity related to decision making and saccade modulation implies that there is a close relationship between the two.

1.5. Saccade vigor modulation

Despite the belief that saccade speed cannot be consciously controlled, in the 21st century, some evidence has been reported that velocity of eye movement might be modulated by visual stimuli. One observation reported eye movements toward images from several categories (Xu-Wilson et al., 2009). It was shown that humans make faster saccades toward faces than neutral objects or random pixels. This result implied that even if saccades were to be made between the same set of positions, their velocity would vary depending on the set of visual stimuli shown at the target location.

To determine whether such effects had anything to do with decision making, a decision-making task was performed where human subjects chose between an immediate reward and a delayed, greater reward while looking at a screen (Reppert et al., 2015). Analysis of data on eye movements showed that not only did saccade vigor decrease after a decision, but saccades were also faster toward the chosen option relative to the other option just before and after the (<1s) decision. The results implied that saccadic behavior is closely related to the decision-making process such that its vigor is faster toward more preferred stimuli.

1.6. Specific aims of thesis

Not only does each human individual have a different baseline level of saccade vigor, but also saccade vigor is modulated within each individual. How are they modulated by properties of visual stimuli in a task where humans make decisions with eye movements? In Chapter 2, we first establish an environment in which the content and location of visual stimuli is associated with reward and effort, respectively, and analyze human subjects' choice of gaze time and saccade vigor from one visual image to another. We show that fixating on a location far away

from the center position requires great effort. We also show that subjects make faster saccades when moving from a high-value image to a high-value image, implying saccade vigor could be affected by richness of environment.

In Chapter 3, we implement the marginal value theorem (MVT), which is a method to explain foraging behavior in environments with localized reward patches. We create a visual foraging task that is analogous to an actual foraging situation where we view images as reward patches and saccades as moving to a new food source. We first show that analysis of gaze time validates MVT in our task environment. Then, we look into how saccade vigor could be modulated by past, present, and future effort, and discuss its significance.

In Chapter 4, we suggest that saccade vigor toward a stimulus could be an indicator of its utility. To validate this, we designed an experiment in which subjects performed a mixture of decision-making trials and cued (forced choice) trials. We show that saccades made in cued trials are faster and quicker toward stimuli with higher utility. We also show that this effect is consistent even when controlled for expected reward across subjects.

With our findings, we suggest two main implications. First, we show that saccadic decision closely resembles the general decision-making process and that saccade vigor is also modulated by factors that affect decision. Second, we show that saccades are faster and quicker toward visual targets associated with higher utility. With these findings, we suggest the possibility of gauging utility with vigor alone, without decisional behavior. This suggestion is intriguing as we show the possibility that utility can be estimated from saccade vigor, a continuous, easily measured variable.

2. Modulation of saccadic behavior by eccentricity and richness of environment

2.1. Introduction

As described in the first chapter, the signal that is transferred to the extraocular muscles (recti) to generate eye movement is composed of a bursting component and an integration component. Since we only consider horizontal saccades, we only take two muscles into account: the medial and lateral recti, which are innervated by oculomotor neurons (cranial nerve III) and abducens neurons (cranial nerve VI), respectively. Neurons in both nuclei, in fact, have been reported to show a discharge rate affected by both eye position and velocity linearly (Robinson, 1970; Van Gisbergen et al., 1981). This implies that perceived motor cost to move the eyes is also dependent on both eye position and velocity. Taking into account that neurons cannot have a negative firing rate, the linear relationships mentioned are actually rectified linear, and since a population of these neurons will innervate the muscles, we could expect that the final signal to the recti would be minimal for center position and zero velocity, being stronger as both the (unsigned) position from the center and saccade speed rises. We can speculate that if fixation on far positions results in great effort, humans will choose to fixate less frequently and for a shorter time on far positions.

In previous studies, it was shown that people move toward faces faster than other types of images (Xu-Wilson et al., 2009). However, what happens if saccadic behavior is directly related to the outcome of a task has not been much explored. In a situation where subjects choose how

long to fixate on and when to move to an image, their decisions will affect the amount of value earned by looking at images. It is this kind of situation that we are interested in.

We generated a task where people would choose how they would look between two images displayed apart for a fixed given time. We are interested in two aspects: (1) Whether the positions of images affect people's behavior of fixation time and saccade vigor, and (2) how vigor is modulated by reward if it is made from one image to another in this task. We propose that high eccentricity will result in shorter fixation, and also play a role in vigor modulation. We also suggest that saccade vigor in this situation is affected by richness of environment (reward of both images).

2.2 Materials and Methods

2.2.1. Participants and general apparatus

Subjects (n=17, 25.2±4.2 years old, mean±SD, 8 females) sat in a well-lit room in front of an LED monitor (59.7 x 33.6 cm, 2560 x 1440 pixels, light gray background, frame rate 144 Hz) placed at a distance of 35 cm. The subject's head was restrained using a bite bar. They viewed images (4x4°, except as noted) and we measured their eye movements using an EyeLink 1000 (SR Research) infrared recording system (sampling rate 1 kHz). Only the right eye was tracked. All subjects were naive to the paradigm. The experiments were approved by the Johns Hopkins University School of Medicine Institutional Review Board, and all subjects signed the written consent form approved by the board. Subjects were paid $15/hour regardless of any behavioral outcome.

2.2.2. Behavioral task

Subjects were presented with a center fixation for 0.5-1 sec, followed by two simultaneously displayed images (3x3°). Each image was selected randomly from five categories: 1-noise, 2-simple shapes, 3-inanimate objects, 4-animate objects, and 5-faces (Fig. 2-1A). Each category consisted of more than 100 images, except simple shapes (45 images). Face images were from those used in a previous study (Liu et al., 2015). Inanimate and animate images were from a standardized data set (Brodeur et al., 2010). We tested all possible combinations of categories, ensuring that each category was presented the same number of times. The two images were always 20° apart, but their positions varied with respect to the midline from 2° to 18°, in 2° increments, chosen randomly from a uniform distribution. Subjects had 2 seconds to freely gaze. During this period the center fixation dot was removed. The trial repeated if the subject spent more than 20% of the total viewing period gazing somewhere other than one of the images. We included trials with two images from the same category twice as frequently as other trials to gather data regarding effects of eccentricity for those specific trials. Subjects completed 9 blocks, and each block consisted of 60 trials.

2.2.3. Data Analysis

The eye position data were filtered with a third-order Savitzky-Golay filter (frame size 11). Saccade onset and offset was determined in real time with 20°/s threshold. We identified saccades between images as only those with onset and endpoint that were within 5° of the boundaries of the start and end images (to account for calibration error). We defined time spent on each image by adding all the times where the gaze was recorded within the image, allowing for microsaccades within the image boundaries.

We used peak velocity as a measure of vigor of saccade. The velocity with which humans move their eyes over a given distance is subject-specific, exhibiting a wide range. Some consistently move their eyes with high velocity, while others consistently move more slowly (Choi et al., 2014; Reppert et al., 2018). Because of this, we focused mostly on the within-subject effect of factors that we tested on vigor. Furthermore, there is an issue of peak velocity varying as a function of saccade amplitude and direction. For saccades from image to image (20° amplitude), we need not account for amplitude. For initial saccades from the fixation dot to the first choice of image, we divided peak velocity by mean within each subject and each amplitude. Thus, for instance, a saccade of 10° amplitude will be normalized by dividing by mean for each subject and for saccades with the same 10° amplitude. We will designate this term as saccade vigor. We will use a more complicated method in our later experiment but will still call those results as vigor too.

To analyze the effects in the experiment, we implemented a linear mixed-effects model that related the dependent variables (gaze duration at the image, saccade peak velocity), to the fixed effect variables (type of image, eccentricity of the image, and type of the other image). Subject label was treated as a random effect. Also, nasal saccades (leftward saccade for the right eye) and temporal saccades (rightward, same case) could have different base levels of vigor, so for velocities, we also had saccade direction as a random effect.

2.3. Results

2.3.1. Gaze time and probability of initial choice reflect preference

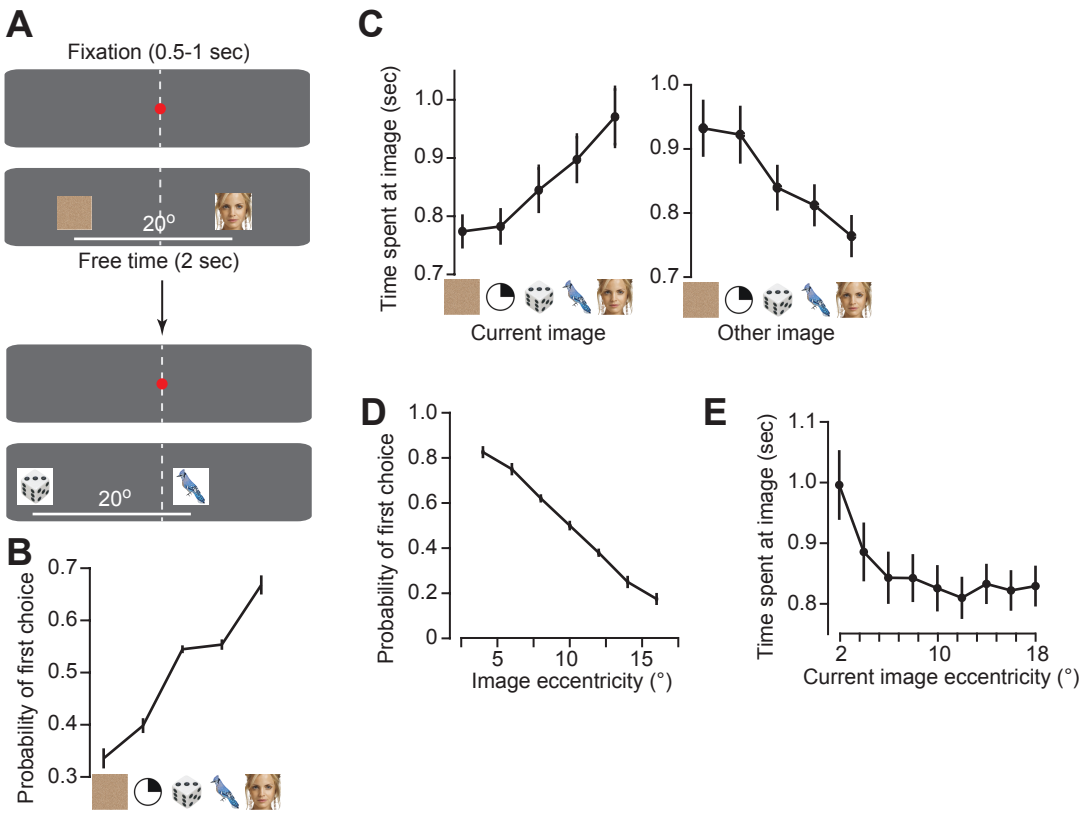

Figure 2-1. Preference is reflected in probability of initial choice and gaze time. (**A**) Experimental protocol. Subjects were asked to freely gaze at either of two images during their display after a brief fixation period. Free time was always 2 seconds. After free time, the images were removed and the fixation dot reappeared, signaling a new trial. (**B**) Probability of initial choice by image category. Images in the x-axis are representative examples of each category, aligned by the order of increasing mean probability. (**C**) Time spent on each type of image (left) and on each type of image on the other side (right). Images in the x-axis are aligned in the same order as in (B) in a monotonically increasing/decreasing order. (**D**) Probability of initial choice by image eccentricity. (**E**) Time spent on current image eccentricity. Difference in time diminishes as eccentricity increases. Error bars for all plots represent mean ± SEM across all subjects.

In this experiment, people could make decisions on how long they would gaze at each image. They would allocate time of 2s between two images in a way that they think gives them the best outcome. Although people show a variety of behavior in each trial, on average they should reflect how much they value each factor. For this task, subjects could choose two things:

how frequently they would choose an image to saccade first toward, and how long they would choose to gaze at an image. Since we controlled all combinations of image categories, the frequency of each image category is independent of the competing image on the other side. Also, we controlled the distance between the images to always be 20° apart, and the probability of an image appearing at each eccentricity was also independent of image category. This control made us analyze the effect of each factor by simply averaging across all trials for each value.

Image category served to modulate reward magnitude. We first established the mean reward of each image category by looking at the probability of first choice. This value increased with image type (Fig. 2-1B, $p < 10^{-15}$), therefore allowing us to establish an order of reward magnitude between image categories. We also looked at mean gaze time for each image category to confirm the reward order. Subjects gazed longer at image categories that they also chose first more often (Fig. 2-1C, left, $p < 10^{-12}$). Also, gaze time decreases as the image that is not looked at gets more rewarding (Fig. 2-1C, right, $p < 10^{-10}$).

We changed the eccentricities of images to modulate gaze fixation effort. We checked if people showed different behavior with images that had different eccentricities. Indeed, subjects showed more frequent first choice toward near center images (Fig. 2-1D, right, $p < 10^{-15}$). Also, subjects gazed longer at images that were nearer to the center (Fig. 2-1E, $p < 10^{-7}$). We could confirm that near center images require less effort to fixate, and thus result in higher preference compared to images far from the center.

2.3.2. Vigor is affected by environmental reward and effort

People make various saccades during this task, including initial saccade from the center dot to their first choice of image, microsaccades while looking at a certain image, etc. We first

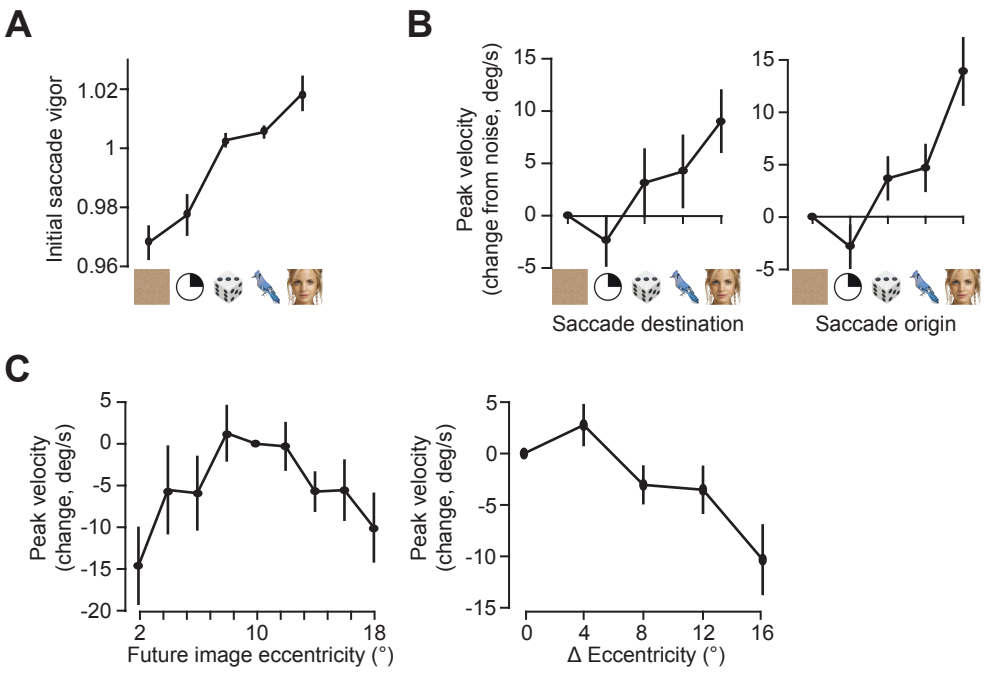

Figure 2-2. Saccade velocity is modulated by reward and effort. (**A**) Initial saccade vigor by image category. Images are aligned in order of preference (see Figure 2-1). Saccade vigor is a normalized value by the main sequence of each subject (see methods). (**B**) Peak velocity of saccades made between images, by image category. Images are always 20° apart. Peak velocity is subtracted by the value for least preferred category (noise). Mean peak velocity is calculated by the image type at destination (left) and the image type at origin (right). Note that both are positive effects. (**C**) Peak velocity of saccades made between images by eccentricity. First, we calculated the mean for each eccentricity value of the target (left). This plot has a U-shape. We also calculated the mean by difference in eccentricities between the target and the origin (right). Error bars for all plots represent mean ± SEM across all subjects.

checked whether we could replicate previous studies that predicted that the initial saccade would be faster for more rewarding stimuli. According to the reward that we inferred from the time and probability measurements, we did find a similar effect (Fig. 2-2A, $p < 10^{-9}$). With this in mind, we particularly analyzed saccades that were made from one image to another. We expected these saccades to be modulated by properties of both images.

We studied the effect of the mean reward magnitude of each image category to see how both images could affect saccade vigor. Obviously, saccade vigor was modulated to be faster toward images that were more preferred (Fig. 2-2B, left, $p < 10^{-8}$), coinciding with previous studies mentioned in Chapter 1. Interestingly, we also observed that saccade vigor was also faster when the previously viewed image was in a more preferred category (Fig. 2-2B, right, $p < 10^{-5}$). Together, we claim that in such a task, saccade vigor is faster when the environment of the trial is richer, disagreeing with the common concept that it is modulated by the difference in reward between the target image and the starting image.

We also tried to analyze the effect of image eccentricity on saccade vigor. We did not see saccade vigor either increasing or decreasing by eccentricity of target image (Fig. 2-2C, left). However, the plot has an inverted U-shape, implying that asymmetry between the eccentricities of two images is the deciding factor. Indeed, we did discover an effect of difference in the eccentricities of two images (Fig. 2-2C, right, $p < 10^{-4}$). We also confirmed that saccade vigor is not faster as the eccentricity of the target image is smaller and the eccentricity of the starting image is larger.

2.4. Discussion

We have proposed a new paradigm on how to observe decision and control of movement by modulating image content and eccentricity. Our proposition is based on the neural circuitry that connects to the extraocular muscles to generate saccades and hold the eyes in the right position. We assumed that since the input signal to the muscles from the cranial nuclei contain both a bursting component and an integrated component, not only moving the eye, but also

holding the eye at a certain eccentricity would require motor effort. Indeed, we saw modulation of gaze time and saccade vigor by eccentricity. Primarily, we saw that people tend to initially choose an image that has a larger eccentricity (i.e., further from the center fixation) more frequently. The subjects are aware that they have enough time to allocate between two stimuli, but they still tend to pick the more preferred one initially, which we observe by modulating image content. The fact that eccentricity decreases both first choice probability and gaze time strongly suggests that both moving the eyes to and holding them in that position require more effort for larger eccentricities.

While the results of first choice probability and gaze time give us compelling evidence that eccentricity modulates effort, our interest mainly resides in our observations regarding saccade vigor, especially the ones made between images. We controlled the distance between images to always be 20° so that we could only compare between saccades with similar amplitude for this purpose. It would have been best if we could have gotten the main sequence for each subject, but since saccades between images have a narrow range of amplitude, we didn't want to risk getting an inaccurate measure of coefficients. Although each subject showed a different baseline level of saccade velocity, the within effect was clearly observed. Interestingly, the effect increases in regard to the reward of both images, instead of the difference between target and starting images. In other words, the richness of the environment decides vigor. Considering that gazing at one image incurs an opportunity cost of not being able to gaze at the other image, this phenomenon is interesting. We could expect saccade vigor to be affected by the difference between two images, since the advantage of gazing at the target is larger if the other target is less preferred. However, we only observed an effect of overall reward. This result led us to consider the marginal value theorem (MVT), which will be explained in more detail in Chapter 3.

We saw a diminishing effect on gaze time as eccentricity gets larger, however, potentially suggesting that the effort function could be nonlinear. This nonlinearity is not expected from the neurophysiological results that extraocular muscles receive an input that has a highly linear relationship with eccentricity (Robinson, 1970; Van Gisbergen et al., 1981). It is also not expected from a simple physical model, if we suppose that muscles are similar to springs holding onto each left/right side of the eyes, which would predict the conservative force (towards the center) to be linear. Some more recent models do take into account acceleration, jerk, or time delay to estimate neural input (Sylvestre and Cullen, 1999), but it is still not fully explainable. Finding out the function of eccentricity in motor effort to hold and move is a potentially very interesting subject.

In regard to this topic, we could also interpret the effect of saccade vigor not decreasing by the difference between target and starting eccentricity. Not only does this mean that saccade vigor is also not dependent on the difference, but on the overall environment, but it also implies that eccentricity modulates motor effort in a nonlinear relationship so that the mean effort to hold on low and high eccentricity is greater than the mean effort to hold on medium values of eccentricity (remember that eccentricity values are paired, i.e., 2° and 18° are always paired, while if one image is at 10°, the other image is also at 10° on the other side, and this happens because we kept the distance between images the same in all trials). This could be because of the average gaze time difference across eccentricity values, since for an eccentricity pair of 2° and 18°, subjects usually fixate on the 2° position longer, resulting in low mean effort to hold compared to other eccentricity pairs. However, we still suggest the possibility of this function being nonlinear.

Since we observed that saccade vigor is decided by starting point property and target point property in the same way, not the difference between the two, it is interesting to develop a new mathematical theory to explain this phenomenon. In fact, we will develop a refined version of the marginal value theorem (MVT) so that we can explain both gaze time (decisional behavior) and saccade vigor (control of movement) with a utility function in Chapter 3.

Chapter 2 is included in our published paper (Yoon et al., 2018).

3. Control of gaze time and vigor in a visual foraging task

3.1. Introduction

Studies have been conducted to explain the decisions that animals make while foraging for food. The model to explain such behavior is called optimal foraging theory (OFT). The idea originated half a century ago (MacArthur and Pianka, 1966), when the authors suggested that animals forage for food in order to maximize net energy gain, thus being in a more advantageous position for natural selection. Ecologists and economists have strived to identify the optimal rule for OFT.

A widely used model was suggested (Charnov, 1976). The author called his model the "marginal value theorem" (MVT). The theory describes how long animals choose to reside in a specific patch of resources (food) that is located discretely, separated by areas that have no food. Animals could choose to not stay and consume food until it is completely depleted, and how long they would stay is a decision-making behavior. Such behavior is predicted when the resources in each patch are limited and it thus results in diminishing returns, so that animals will get less food per time as the patch gets depleted. The founder of the theorem tried to estimate the time, which is affected by properties of the patch and also by environmental factors.

The original paper describes a situation where animals try to maximize the net energy intake rate by choosing the time spent in each patch. With a series of simple calculations, it concluded that the optimal time the animal would choose to spend in each patch to maximize net energy intake rate is when the marginal capture rate (derivative of energy assimilation function)

drops to the average capture rate for the environment. In other words, the animal would choose to leave when their current intake rate drops to the average intake rate, so that staying longer is less beneficial than the average gain. While this theorem explains well how long animals would stay, it does not explain how quickly they would travel between patches, because the theorem assumes that interpatch travel time is known and travel energy cost is not affected by other factors.

We developed a new version of the marginal value theorem by introducing travel energy cost as a function of interpatch travel time. We could explain how movement vigor could be affected by patch properties and environmental factors. We also conducted an experiment that introduced a foraging-like situation. We discovered that gaze time on each image agreed with our predictions from the marginal value theorem. Also, we observed modulation of saccade vigor by several factors, which also justified our new modified MVT model.

3.2. Materials and Methods

3.2.1. Generalized MVT

Our generalized MVT model is very similar to the original version but we take into account the cost of movement varying by travel time as well. In this case, subjects could control two variables: duration of time spent harvesting in each patch, and travel time to that patch. First, we establish the reward acquired and effort expended at patch $n = 1, \dots, N$. We make an assumption that reward accumulation is an increasing function but then saturates, while consistently spending motor effort. We represent the harvest function $f^{(n)}$ as the sum of reward acquired and effort expended during time spent:

$$f^{(n)}\left(t_h^{(n)}\right) = \alpha^{(n)}\left(1 - \frac{1}{1+\beta t_h^{(n)}}\right) - u_h^{(n)}\left(t_h^{(n)}\right) \tag{3.1}$$

In this formula, $\alpha^{(n)}$ is the total amount of reward in the patch, $t_h^{(n)}$ is the harvest time, β is rate with which the reward is harvested, and $u_h^{(n)}$ is the effort required for harvesting (also a function of harvest time). The exact form of this function is not critical here. What is important is that the harvest function is increasing but has a diminishing rate of return, with its second derivative being negative (Fig. 3-1A). We assume that harvesting effort has a linear relationship by time:

$$u_h^{(n)}\left(t_h^{(n)}\right) = k^{(n)} t_h^{(n)} \tag{3.2}$$

In this formula, $k^{(n)}$ is the rate of harvest effort by time. Realistically, all variables are positive, so we got the harvest function that satisfied our assumption.

We can also develop the local capture rate $J^{(n)}$ in each patch, also taking into account the effort expended to travel to the patch ($u_m^{(n)}$), and travel time to patch $t_m^{(n)}$. The traveling effort $u_m^{(n)}$ will also depend on the travel distance $d^{(n)}$. Then we have the following:

$$J^{(n)} = \frac{f^{(n)}\left(t_h^{(n)}\right) - u_m^{(n)}\left(d^{(n)}, t_m^{(n)}\right)}{t_h^{(n)} + t_m^{(n)}} \tag{3.3}$$

As the original version, the objective would be to maximize global capture rate $\bar{J}$:

$$\bar{J} = \frac{\sum_{n=1}^{N} f^{(n)}\left(t_h^{(n)}\right) - u_m^{(n)}\left(d^{(n)}, t_m^{(n)}\right)}{\sum_{n=1}^{N} t_h^{(n)} + t_m^{(n)}} \tag{3.4}$$

We didn't specify the function for $u_m^{(n)}$ yet, but we assume that this function is convex, with its second derivative positive (Fig 3-1F). We think this is justified, because energy expenditure in many types of movements is convex, including walking and reaching (Zarrugh et

al., 1974; Shadmehr et al., 2016). This assumption is critical, because only then should we be able to derive the expression for optimal movement duration.

Suppose that the subject considers moving to and harvesting in patch n. Then, we can only attend to the terms that are related to this patch:

$$\bar{J} = \frac{f^{(n)}\left(t_h^{(n)}\right) - u_m^{(n)}\left(d^{(n)}, t_m^{(n)}\right) + A}{t_h^{(n)} + t_m^{(n)} + B} \tag{3.5}$$

Constants A and B contains all other terms that are not associated with behavior in this patch. Subjects control harvest time $t_h^{(n)}$ and travel time $t_m^{(n)}$ to maximize the above. To get the optimal values, we note that the derivatives of $\bar{J}$ with respect to the two time periods can be simply written as:

$$\frac{d\bar{J}}{dt_h^{(n)}} = \left(\frac{df^{(n)}}{dt_h^{(n)}} - \bar{J}\right)\frac{1}{t_h^{(n)} + t_m^{(n)} + B} \tag{3.6}$$

$$\frac{d\bar{J}}{dt_m^{(n)}} = \left(\frac{du_m^{(n)}}{dt_m^{(n)}} + \bar{J}\right)\frac{-1}{t_h^{(n)} + t_m^{(n)} + B} \tag{3.7}$$

The optimal values of the two variables are found when the two derivatives are simultaneously equal to zero. As a result, the optimal harvest duration will be specified as the relationship between the harvest function and the global capture rate:

$$\left.\frac{df^{(n)}}{dt_h^{(n)}}\right|_{t_h^{(n)*}} = \bar{J} \tag{3.8}$$

Also, the optimal movement duration will be specified as the relationship between the effort expenditure function and the global capture rate:

$$\left.\frac{du_m^{(n)}}{dt_m^{(n)}}\right|_{t_m^{(n)*}} = -\bar{J} \tag{3.9}$$

Asterisks designate the optimal values of each variable. The expression for harvest function is still similar to the original version of the marginal value theorem, although now global capture rate $\bar{J}$ is also a function of travel time. In addition, we now have a new expression for optimal travel time.

The meaning of these equations is fairly intuitive. For the harvest time, the optimal time is when the "marginal rate" drops to the global capture rate. For the travel time, the optimal duration is when the "marginal gain" (how much less effort is spent as movement time increases) drops to the global capture rate (minus sign for a negative slope). The two equations reach similar conclusions.

3.2.2. Generalized MVT and its applications

To test the model, we should be able to build an environment that satisfies basic assumptions for the model. We create an environment where subjects are given patches of images on a screen. They are free to move between patches as they want to. In this case, they can decide both their own gaze duration on each image and saccade duration between images. We control reward $\alpha^{(n)}$ via image type and effort required to harvest $u_h^{(n)}$ via image eccentricity. This assumption holds valid because of our previous observations in Chapter 2: People choose more often and gaze longer at certain image types and images that have low eccentricity (Fig. 2-1). The gain by looking at an image would diminish as subjects become more familiar with that image. The effort to hold would still remain over time. Although we cannot numerically derive reward values, we can still predict how each factor would affect gaze time and saccade duration.

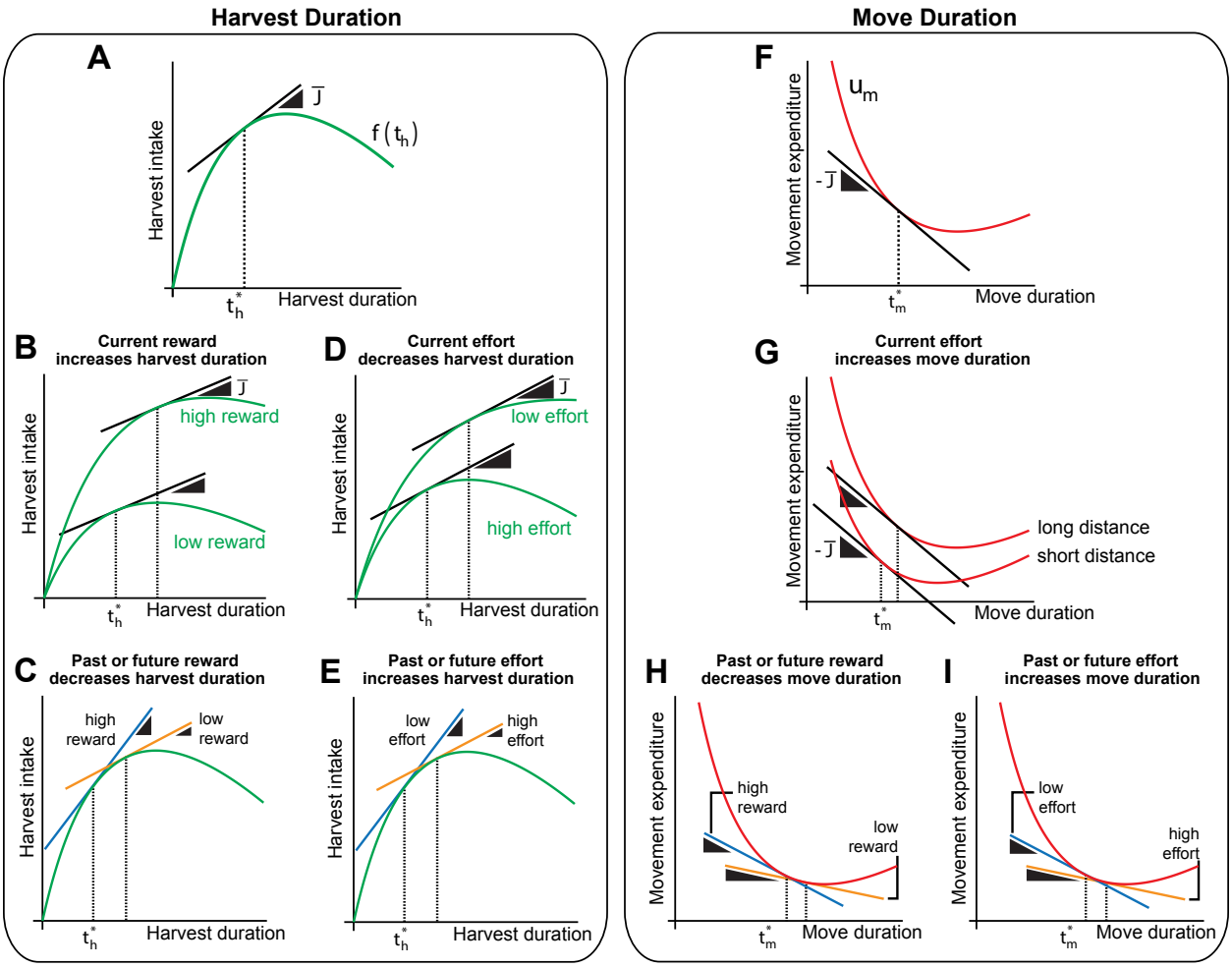

Figure 3-1. The generalized marginal value theorem predicts harvest duration and movement vigor during foraging. (**A**) The harvest function is the sum of accumulation of reward with a diminishing rate and expenditure of effort with a constant rate, resulting in its concave form. For such harvest functions, the theory predicts the individual to leave the patch when the marginal rate equals the global rate. (**B**) Increased reward result in longer harvest duration for the same global rate. (**C**) Similarly, increased effort results in shorter harvest duration. (**D**) When the harvest function stays the same, increased past or future reward would change the global rate, thus changing the optimal harvest duration. (**E**) Similarly, past or future effort would have such effects. (**F**) During movement, for an effort function that is convex by movement duration, the theory predicts the individual to choose vigor so that the marginal advantage of movement duration equals the global rate. (**G**) The theory explains why longer distance is required to move results in longer movement duration. This does not imply that the speed is slower, because individuals will move different distances. (**H**) Although it has no direct relationship with movement, past or future reward is predicted to change movement vigor by modulating the global rate. (**I**) Similarly, past or future efforts will have their own effects.

For example, the theorem predicts the reward of the current image to increase harvest

duration (Fig. 3-1B) because it delays when the marginal rate drops below the global capture rate. Similarly, it predicts the eccentricity of the current image (=holding effort) to decrease harvest duration (Fig. 3-1D) because it does the opposite. However, not only does the modulation of the current harvest function change the optimal gaze time, but change in the global capture rate does as well. This would include both past and future reward and effort. An increase in reward would increase the global capture rate and thus decrease the optimal harvest time (Fig. 3-1C) (in other words, richness of environment results in a shorter gaze time with the current image). Similarly, an increase in effort would decrease the global capture rate and thus increase optimal harvest time (Fig. 3-1E) (in other words, hardness of environment results in longer gaze time with the current image).

In addition, we could also predict how several factors would affect saccade vigor. We can even explain why longer distance results in a longer saccade duration (Fig. 3-1G): change in effort function results in delay of when marginal gain would drop below the global capture rate. Of course, modulation of the global capture rate also has an effect on saccade vigor: an increase in reward and decrease in effort would increase the global capture rate and thus decrease optimal movement duration, resulting in a faster saccade for the same amplitude (Fig. 3-1H, 3-1I). We will try to show that such modulations indeed appear in our results.

Interestingly, one alternative way to define the global capture rate is to derive it as an average value of the individual utilities of each patch:

$$\bar{J} = \frac{1}{N}\sum_{n=1}^{N} J^{(n)} = \frac{1}{N}\sum_{n=1}^{N} \frac{f^{(n)}\left(t_h^{(n)}\right) - u_m^{(n)}\left(d^{(n)}, t_m^{(n)}\right)}{t_h^{(n)} + t_m^{(n)}} \tag{3.10}$$

In this case, the optimal harvest duration and movement time depend only on the current conditions – that is, past and future reward and effort (and anything that only affects the global

capture rate) should not affect decision making and motor control. We will set out to test these predictions and prove that it is not the case.

3.2.3. Participants and general apparatus

Subjects (Experiment 1: n=22, 30.0±11.4 years old, mean±SD, 10 females; Experiment 2: n=18, 23.1±5.2 years old, mean±SD, 9 females) sat in a well-lit room in front of an LED monitor (59.7 x 33.6 cm, 2560 x 1440 pixels, light gray background, frame rate 144 Hz) placed at a distance of 35 cm. The subject's head was restrained using a bite bar. They viewed images (4x4°, except as noted) and we measured their eye movements using an EyeLink 1000 (SR Research) infrared recording system (sampling rate 1 kHz). Only the right eye was tracked. All subjects were naive to the paradigm. The experiments were approved by the Johns Hopkins University School of Medicine Institutional Review Board, and all subjects signed written consent form approved by the board. Subjects were paid $15/hour regardless of any behavioral outcome.

3.2.4. Behavioral task – Experiment 1

Subjects began each block of trials with an image (4x4°) at the center location and a dot at another position along the horizontal axis. The dot location was selected randomly (uniform distribution) at a distance of 10°, 15°, or 20° from the image along the horizontal axis, but never beyond 20° from the center of the screen. We employed three image types: 1-simple shapes, 2-realistic objects, and 3-faces, with at least 500 images in each category. Once the gaze shifted to the dot, the dot disappeared, and an image appeared at its location (Fig. 3-2A). In each trial, the image category was selected randomly with equal probability from each category, and then with equal probability from within that category. As there were no constraints on duration of gaze,

block length was defined in time (200s) and not trials. Behavior was examined during 10 blocks, resulting in an average of 2812±419 trials per subject.

3.2.5. Behavioral task – Experiment 2

The basic experiment design was similar to Exp. 4 in that the subjects viewed an image for a duration of their choosing, being given the option of making a saccade to a dot, after which they were presented with an image at that location. Subjects experienced 16 blocks, 100 trials/block. To modulate effort history, during a block, all images (except probe images) appeared at high eccentricity in a range from 10° to 30° with respect to midline (uniform distribution, mean eccentricity of 20°). In another block, the images appeared at low eccentricity within 10° of the midline (mean eccentricity of 5°). These blocks represented high effort and low effort environments, respectively (Fig. 3-3A). In each block, 20% of the trials were a probe, placing the image at 10° with respect to the midline, and the dot on the contralateral side, also at 10° with respect to midline. Every fifth trial was a probe trial. Images were selected randomly from three categories, the same as in experiment 1. However, every probe trial started with an image from category 2 (realistic object).

3.2.6. Data Analysis

The eye position data were filtered with a third-order Savitzky-Golay filter (frame size 11). Saccade onset and offset was determined in real time with 20°/s threshold. We identified saccades between images as only those with onset and endpoint that were within 5° of the boundaries of the start and end images (to account for calibration error). We defined the time

spent on each image by adding all the times where the gaze was recorded within the image, allowing for microsaccades within the image boundaries.

We used peak velocity as a measure of vigor of saccade. The velocity with which humans move their eyes over a given distance is subject-specific, exhibiting a wide range. Some consistently move their eyes with high velocity, while others consistently move more slowly (Choi et al., 2014; Reppert et al., 2018). Furthermore, peak velocity varies as a function of saccade amplitude and direction. To quantify the effects of reward and effort on the saccade of each subject, we first built a subject-specific model of peak velocity as a function of amplitude and direction using a maximum likelihood approach, as described in another study (Reppert et al., 2018). The data that we used for constructing the model consisted of all saccades made by the subject during the experiment. Given the measured amplitude and direction of a given saccade, the model predicted the expected saccade peak velocity for that subject. We defined saccade vigor as the ratio of the actually measured peak velocity with respect to the predicted velocity. This normalization would account for vigor modulation accounting for individual difference and the main sequence (Section 1.1).

To analyze the effects in the experiment, we implemented a linear mixed-effects model that related the dependent variables (gaze duration at the image, saccade peak velocity), to the fixed effect variables (type of image, eccentricity of the image). Subject label was treated as a random effect. In cases where we compared “high” and “low” values (two categories), the linear model is same as repeated measures ANOVA.

3.3. Results

3.3.1. Effect of reward on gaze duration

This experiment created an environment that is analogous to a foraging situation where subjects gazed at each image for a certain amount of time (analogous to harvesting food from a patch), and then moved to the next image (analogous to moving to a new patch of food). Thus, if the model is valid, we should be able to see the predictions made by the model (Fig. 3-1). Subjects could consciously choose the gaze duration for each image, and according to our version of the MVT model, modulate saccade velocity moving between images. Thus, we expected changes in these two parameters by various local and global factors to be different across trials.

In Chapter 2, we already confirmed that image categories alter gazing behavior. In other words, subjects look longer at the more preferred image. Our results also show that this is true. Subjects certainly spend a different amount of time on each image category, on average (Fig. 3-2B, $p < 10^{-5}$). The order of gaze time is identical to what we already saw in Chapter 2, although for this experiment there was no time limit and no competing image. Our model did predict the current reward to increase gaze duration by modifying the harvest function (Fig. 3-1B).

In our task, there is no way for the subjects to predict what image type would appear in the next location, where the dot is, so future reward would not have any effect. However, an already experienced reward is predicted to have an effect by changing the global capture rate (Fig. 3-1C). We did see an effect of image type at just the previous patch (Fig. 3-2C, $p < 10^{-3}$), toward the opposite direction, which was predicted by the model. This effect was very small, however, with a range only roughly one-tenth of the effect of current image type. Due to the large variance in subjects' behavior, we did not see any effect of image type past one previous patch.

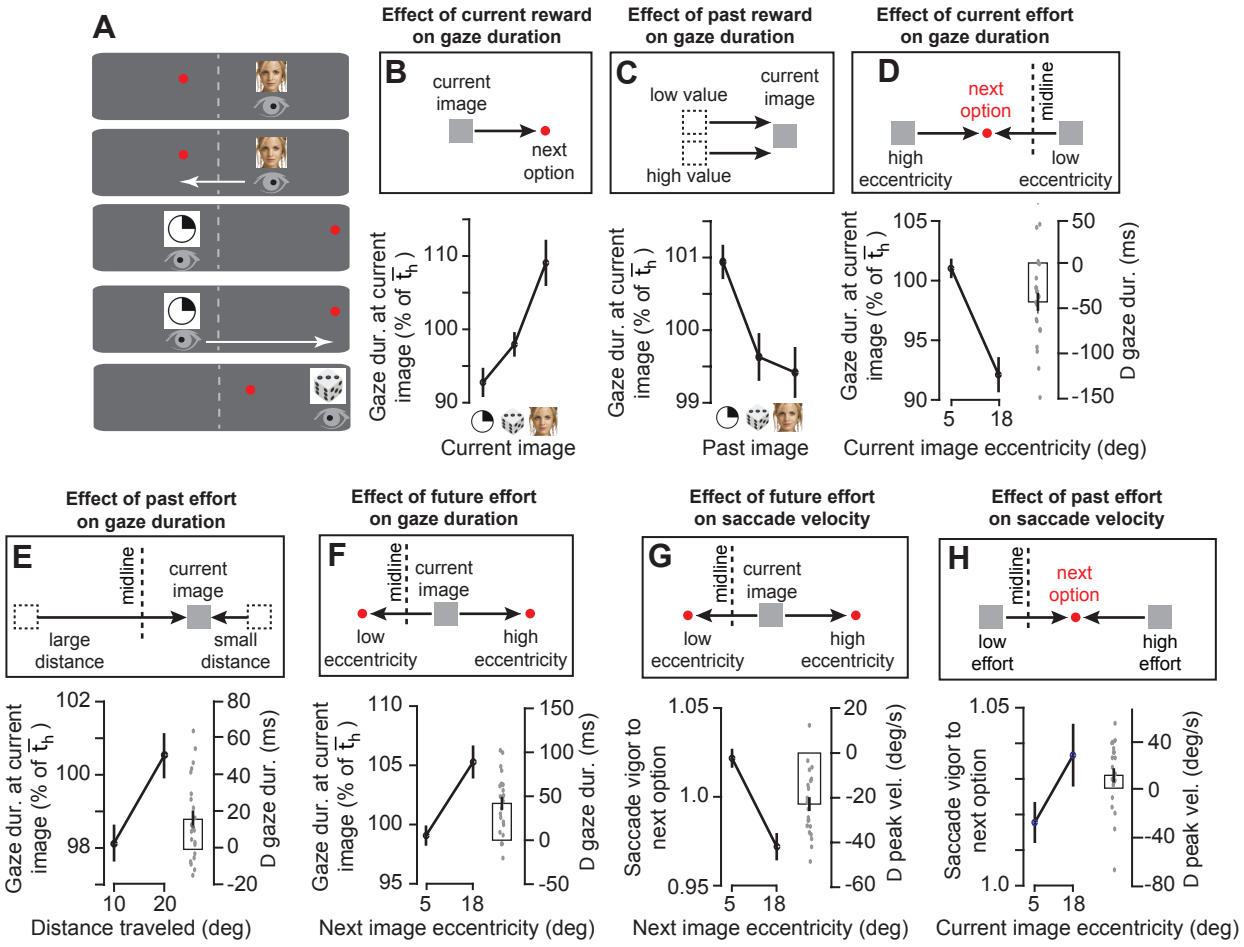

Figure 3-2. Various factors modulate gaze duration and saccade vigor in a visual foraging task. (**A**) Experimental protocol for experiment 1. Subjects make continuous decisions of whether to continue gazing, or to move their eyes to the next target. (**B**) Gaze duration by current image category. (**C**) Gaze duration by past image category. Past reward decreases gaze duration. (**D**) Gaze duration by current eccentricity. Higher eccentricity requires more effort during holding, therefore changing the form of the harvest function. (**E**) Gaze duration by past distance moved. Past effort modulates the global capture rate. (**F**) Gaze duration by future eccentricity. Future effort also modulates the (expected) global capture rate. All gaze durations are normalized for each subject by mean across all trials. (**G**) Saccade vigor by target image eccentricity. A similar principle as in (F). (**H**) Saccade vigor by current image eccentricity. The effect is the opposite of the effect of target image eccentricity. Saccade vigor is calculated by normalization using each subject's main sequence (see methods). Error bars for all plots represent mean ± SEM across all subjects.

3.3.2. Effect of effort on gaze duration

Although the task instructions were simple and many variables for eccentricity were randomly chosen, parameters of eccentricities and distances were moderately correlated because of some task conditions we applied and because distances were calculated as differences between signed eccentricities, thus being completely dependent. Therefore, we needed to adopt an additional method to single out each parameter to study its effect on subjects' behavior. To do this, we simply extracted pairs of sets of parameters that only differed by our choice of the parameter of interest, while keeping the other parameters that were correlated to be the same. This simple procedure excluded many trials for each test, but still provided us with enough statistical power because we had many trials to begin with (see methods). In this way we were able to test the individual effects of each parameter without having to be concerned with cross-correlation between independent variables.

As mentioned earlier in Chapter 2, image eccentricity modulates behavior by modulating the effort to fixate on the right location. We observed that eccentricity of the current image had a negative effect on gaze duration (Fig. 3-2D, $p < 10^{-10}$). Our model predicted that this would happen because eccentricity would modify the value of $k^{(n)}$, the coefficient for holding effort, which results in a change of the harvest function (Fig. 3-1D).

Not only did eccentricity of the current image modulate gaze duration, but also those of past and future images, as well as distances between images. We also observed that distance traveled before fixation on the current image had a positive effect (Fig. 3-2E, $p = 0.046$). In other words, subjects gazed longer at a target image just after they had to move their eyes a longer distance. This effect is predicted from our model because distance traveled would decrease the global capture rate, thus increasing gaze duration on the current image (Fig. 3-1E).

Similarly, the eccentricity of the future image had a similar effect (Fig. 3-2F, $p < 10^{-12}$). This effect is also predicted from a decrease in the global capture rate, same as the distance from the past image (Fig. 3-1E).

3.3.3. Effect of effort on vigor

For this task, we are also interested in observing the effects of eccentricity and distance on vigor. Since distance itself would modulate the baseline level of velocity due to the main sequence of saccade, the effect of distance traveled would be trivial. Thus, we investigated the effect of eccentricities on saccade vigor. Unlike gaze duration, there is no concept of "current" image, so all eccentricities will be past or future. According to our model, both are predicted to lower the global capture rate and thus increase saccade duration (= decrease saccade vigor) (Fig. 3-1I). However, we observed some conflicting results.

To compare between different sets of parameters that differ in saccade distance, we used a normalization method mentioned in the methods section (Section 3.2.6) (Reppert et al., 2018). We observed that the normalized vigor is affected by both eccentricities; it decreases by eccentricity of target image (Fig. 3-2G, $p < 10^{-4}$) and increases by eccentricity of current image (Fig. 3-2H, $p < 10^{-15}$). Only the former effect is expected by our model, as illustrated (Fig. 3-1I).

In conclusion, we observed effects of past and future efforts on gaze duration and saccade vigor, and they mostly agreed with the predictions from our model. However, we fond one mismatch where past effort expenditure resulted in higher vigor. The next experiment is to further confirm this observation in this task and provide an alternative explanation of how vigor could be modulated by past effort.

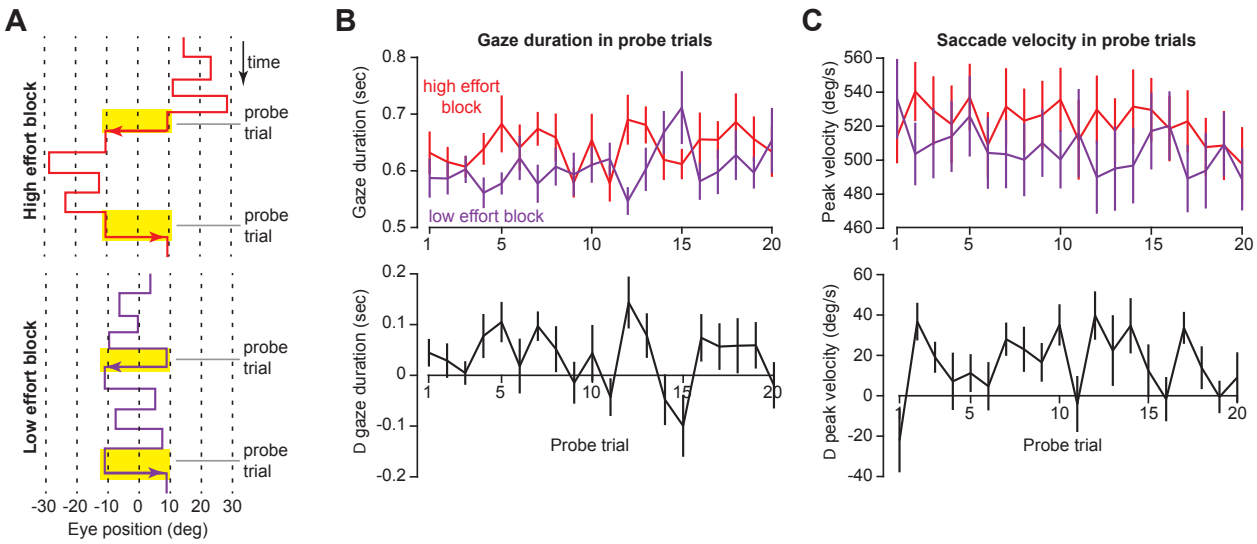

Figure 3-3. History of effort causes an increase in movement vigor. (**A**) Experimental protocol for experiment 2. The procedure is identical to experiment 1, but in this experiment, subjects experienced a "high effort" block and a "low effort" block that are clearly distinctive. Effort is modulated by history of eccentricities in the images displayed. (**B**) Gaze duration for each n-th probe trial. Plots show separately the gaze duration for different blocks (top) and the difference between the two quantities (bottom). (**C**) Saccade peak velocity for each n-th probe trial. Plots are drawn similarly to (B). Error bars for all plots represent mean ± SEM across all subjects.

3.3.4. Effect of past effort expenditure on vigor

Experiment 2 in this section was conducted to show two things: that (1) the effect of past effort expenditure on vigor that we have observed (Fig. 3-2H) also happens in the relatively long term, and (2) the effect is observed even when saccade trajectories are controlled to be the same (i.e., saccades that have the same starting position, target position, and amplitude). In order to show these, we conducted a task similar to experiment 1 but controlled parameter distributions so that they have distinct "environments" (Fig. 3-3A). In a high effort environment, targets were presented in positions that were at least 10° eccentricity. In a low effort environment, targets were presented in positions that were at most 10° eccentricity. Meanwhile, we had some probe

trials in each environment that had its current image at 10° and the next target dot at 10° on the other side, so that it satisfied both environment conditions. This manipulation allowed us to compare between probe trials that have the same local properties but in different environments.

Similar to experiment 1 in this chapter, we still see longer gaze duration in probe trials in the higher effort environment (Fig. 3-3B, $p < 10^{-4}$). Furthermore, we also see faster saccade velocity for probe trials in the higher effort environment (Fig. 3-3C, $p < 10^{-9}$). The fact that there still is saccade vigor modulation makes us believe that past and future efforts (represented by eccentricities) to fixate on images modulate such that higher effort results in faster saccade vigor. This is also still equivocal with results in experiment 1.

3.4. Discussion

We proposed a new model in the context of foraging and suggested a way to explain both time in patch (= gaze duration on image) and travel time (= saccade vigor, reciprocal). Previous models simply assumed a certain value of mean travel time and used it to explain duration in patch, under the assumption that travel time is unaffected; we proposed that even movement vigor could be modulated as a variable. We also suggested that a foraging situation could be formulated by presenting visual targets in various positions, which we showed represent harvesting effort (= holding effort) in Chapter 2. In fact, this context of study has been conducted for a long time under a task termed "visual search," but we specifically confine variables in order to observe and quantify systematic changes in gaze duration and saccade vigor.

Our generalized theory of the marginal value theorem (MVT) states how gaze duration and saccade vigor could be modulated not only by local reward and effort, but also by average

global capture rate. The predictions from our model come from the logical policy to maximize global utility, defined as the sum of all rewards divided by the sum of all time spent. Although there could be some modifications to this formula, such as adopting weights (<1) that is multiplied as it gets further away toward the past/future from the current trial of interest, as is done for temporal difference (TD) learning (Sutton, 1988), the conclusion should still remain similar and predictions should be affected in a similar way by local and global factors. We do claim that our definition of global utility is important considering that an alternative hypothesis of defining global utility as the mean of utility of each trial predicts the global capture rate to have no effect on current gaze duration, which is a completely different prediction.

In this chapter (and Chapter 2 as well), we used an analogy of our task to an actual foraging situation. We assumed that viewing images was akin to harvesting food on a patch. This comparison comes from behavioral evidence (Xu-Wilson et al., 2009) and neurophysiological evidence that people like and move faster toward faces, and that facial images trigger a response in regions associated with reward (O'Doherty et al., 2003). As with facial images, we could assume that images in other categories would show similar effects to varying degrees. In addition, we assumed that making a saccade from one image to another is analogous to traveling between harvest patches, and that while looking at an image, the eye exerts effort similar to how animals have to spend effort to consistently harvest on a patch. As mentioned in Chapter 2, eccentricity results in both the neurophysiological consequence of a higher firing rate for motor neurons and the behavioral consequence of fixating for a shorter time due to requiring more effort. Thus, we could safely make the analogy, and apply our model to predict saccadic behavior.

The pattern of gaze duration on each image spent by subjects showed exactly the pattern predicted by our new version of the marginal value theorem. We showed that current reward lengthened gaze duration, obviously. We also observed that a reward (image type) in a previous trial shortened gaze duration, although its effect size was much smaller than the effect of the current trial reward. The fact that no more previous reward showed any significant effect indirectly points to the fact that there should be a temporal discount term, as mentioned above, even though this story alone does not reveal what its exact mathematical form would be. Nevertheless, the effect of past reward implies that a decreased global capture rate results in a longer gaze duration on the present image.

Also, we observed that images on higher eccentricities are gazed at for a shorter time, implying that the eye has to exert consistent effort during gaze time. Not only this, but we also saw that past and future efforts (represented by distances and eccentricities) affect gaze duration, but in the opposite direction (so that they lengthen gaze duration in the present trial). Our model predicted this would happen because of alteration of the global capture rate, and the predictions match the results. These effects also trail beyond one trial back and forward, also implying a temporal discount. In fact, temporal discount of reward had been observed in saccadic experiments (Shadmehr et al., 2010). With these observations, we believe that our analogy between our experiments and foraging is reasonable and our models are applicable to the situation of our experiments.

Since our global utility rate term is not simply the average (or weighted sums, if we consider time decay) of utilities of all trials, the question would be how a temporal discount term could be applied to the formula. This cannot be shown simply from our study, but we could simply suggest the discount multiplier in each trial term for both reward in the nominator and

time in the denominator. Even this simple assumption would predict that the effects of previous trials would decay as it is further away in the timeline toward the past, as a change in the parameters of that trial would have much less of an effect on the global rate. It would still be an interesting topic to look at how the global rate is actually affected by each trial based on the subjects' foraging behavior.

Recently, a number of studies have shown that movement vigor is modulated by reward, including saccadic eye movements (Takikawa et al., 2002; Xu-Wilson et al., 2009; Summerside et al., 2018). We observed effects of past and future efforts on saccade vigor, suggesting that this is affected by utility and not just reward. In fact, some of these effects have been shown in previous studies (Haith et al., 2012). However, the observations do not exactly match the predictions from our model. Our model predicted that vigor would decrease because the global capture rate would drop, but it did not happen for past effort, since we observed it actually resulting in higher vigor.

One suggestion to explain this phenomenon is simply that a different saccade trajectory is what is causing the difference in vigor. It has been suggested that saccades of different directions and relative positions could have different peak velocities, even if they have the same amplitude, such as left-right asymmetry (Vergilino-Perez et al, 2012) and nasotemporal asymmetry (Jóhannesson and Kristjánsson, 2013). Nonetheless, we did try to minimize this effect by accounting for both directions (leftward and rightward). Previous studies do not show systematic differences in saccade velocity by its starting and ending eccentricities, so we believe that our observations are due to the structure of the task requiring holding on each image before and after the eyes arrive. This is further confirmed by experiment 2, where we saw that this effect occurs even when we compare between saccades that have the same starting and ending eccentricities.

Not only does this result show that long-term effort has an effect on current vigor, but it also shows that even when saccade trajectories are controlled, the vigor is modulated by effort, thus further confirming our suggestion.

An alternative way to explain this phenomenon is to relate it to a similar effect. In fact, people elevate the subjective value of a reward when it is gained by an expenditure of cost. This is usually referred to as effort justification (Aronson and Mills, 1959). Although this behavior was originally a psychological phenomenon, similar behavioral patterns were observed in animals foraging (Clement et al., 2000; Kacelnik and Marsh, 2002). It has been suggested recently that the reason for this phenomenon is a within-trial contrast (Zentall, 2013). In detail, it explains that since a reward associated with higher expenditure of effort used to have low utility, when it is evaluated without concerning any cost, its value elevates due to the contrast effect. In economics, a similar effect is referred to as the sunk cost fallacy, where people value things they have obtained by paying more for them.

Recently, a series of neuroscientific studies suggested the plausibility of this phenomenon (although it was not directly addressed). A study that observed the behavior of mice in distinct types of environments that differed by whether the cost to obtain food increased/decreased showed that dopamine neurons fired stronger in an environment where cost decreased under the same cost value, indicating that they had a higher firing rate succeeding high-effort expenditure (Schelp et al., 2017). Another study that looked at the relationship between dopamine signals and movement showed that substantia nigra pars compacta (SNc) dopamine activation promotes succeeding movement initiation (da Silva et al., 2018). These facts considered together could suggest that higher expenditure of effort would lead to more vigorous movement, which we saw in our observation.

Neural patterns involving foraging behavior have already been observed in the frontal lobe. When fixating, animals are in a persisting decision process. They have to make a decision to either stay or leave the current patch. The effort of eye fixation is encoded in the frontal eye field (FEF) (Segraves, 1992), while the relative value of leaving is encoded in the cingulate cortex (Hayden et al., 2011), whose rate rises faster when the effort expenditure of travel is larger. When this rate reaches a threshold, the animal would decide to travel to the next patch. The cingulate cortex is also affected by reward, as its rise is faster when the current reward rate is higher, and the threshold is also affected by the reward history (Barack et al., 2017). Together, we see the similarity between our behavioral observations and the pattern of neuronal activity in the frontal lobe.

As already mentioned in Chapter 1, dopamine neurons in the SNc innervate the basal ganglia, while neurons in substantia nigra pars reticulata (SNr), part of the basal ganglia, send inhibitory output to the superior colliculus. In fact, a bilateral lesion of the globus pallidus externus (GPe), which receives inputs from other areas of basal ganglia and sends output to the SNr, results in the inability of animals to modulate saccade vigor in response to changes in reward (Tachibana and Hikosaka, 2012). Thus, it seems clear that dopamine release is crucial in modulating saccade vigor given a decision-making task.

However, phasic dopamine release only reveals an effect of local reward and effort. It has been suggested that tonic dopamine release could encode a sustained reward history (Niv et al., 2007). Unfortunately, not much has been studied about this, although several effects of tonic dopamine release have been reported (Beeler et al., 2010). Other studies mention that reward history is instead reflected in tonic firing of serotonergic neurons (Cohen et al., 2015). Thus,

while it is clear that the history of reward and effort modulates vigor, its neural basis is poorly understood.

In this chapter (and also Chapter 2), we use images that are at least neutral (some of them, such as faces, could even be considered attractive). In the case of an aversive image, vigor could be modulated differently. It has been reported that appetitive and aversive outcomes incur different patterns of neuronal activity in the prefrontal cortex (Kobayashi et al., 2006) and dopaminergic neurons (Matsumoto and Hikosaka, 2009). In fact, serotonergic neurons in the dorsal raphe nucleus (DRN) also show distinct firing patterns in both types of outcomes (Hayashi et al., 2015). Meanwhile, the effect of positive or negative images has mostly been studied in the field of psychology in terms of emotion and attention. In fact, it has been observed that saccades occur more quickly towards more fearful tasks (Öhman et al., 2001, Bannerman et al., 2009). It would be interesting to see how movement vigor differed in the presence of negative stimuli.

Our experiments had the great advantage of being able to collect thousands of trials of fixation and movement of the eyes in a relatively short time of about 1 hour. Also, we used images to provide accumulating reward during fixation, and this assumption seems valid given our observations. In fact, we did have to make important assumptions about the harvest function and the cost of travel function; in order to apply our model, the harvest function should be concave, and the movement effort function should be convex. These assumptions will need to be tested in a more quantified study where reward value and cost of movement is objectively measurable. Also, it could be a concern that the time scale is very different from the situation that the original MVT usually explains, but since predictions turn out to be valid, we could still make the conclusion that people still care about optimizing global capture rate in those situations, even if the effort required is very small and the time scales are very short.

Chapter 3 is included in our published paper (Yoon et al., 2018).

4. Modulation of vigor and reaction time to visual stimulus associated with reward

4.1. Introduction

Subjective value is a popular topic in economics, and is usually referred to as utility, which is defined as the amount of worth or value gained by the option of choice. This value is measured from people's decision and choice behavior. Most studies involve "risky" choices, which present various outcomes with differing probability. It was initially suggested that people choose the option with the highest expected utility value (von Neumann and Morgenstern, 1944). A series of studies led by economists eventually arrived at the Prospect theory (Kahneman and Tversky, 1979), which suggests that humans make decisions based on their own heuristics on potential values of gains and losses, which accounts for some decisions that do not follow the expected utility theory (Allais, 1953; Ellsberg, 1961). Thus, not only is utility an abstract quantity that cannot be measured directly, it is also subject to subjective heuristics and thus needs carefulness in measurement.

A number of studies revealed that movement vigor is modulated by reward and effort (Yoon et al., 2018, Shadmehr et al., 2019, Summerside et al., 2018). In fact, recent studies have even shown that reward prediction error (RPE) modulates movement vigor in saccade adaptation (Sedaghat-Nejad et al., 2019). Interestingly, RPE is also what is encoded by dopaminergic neurons in the substantia nigra pars compacta (SNc) (Schultz et al., 1997; Bayer and Glimcher, 2005). It was also discovered recently that stimulation of dopamine just before movement increases its vigor (da Silva et al., 2018). This series of results raises the possibility that the vigor

with which an individual moves toward an option depends on its subjective value assigned by the brain. If this is the case, vigor could provide an objective assay of subjective value.

Alternatively, instead of representing subjective value, it might reflect the salience of the stimulus. Chapter 3 already mentioned that fearful images are responded to more quickly (Öhman et al., 2001, Bannerman et al., 2009). If this is the case, vigor then increases for both larger gain and loss. Some dopamine neurons also show increased activity for both positive and negative feedback (Matsumoto and Hikosaka, 2009). It was already mentioned in Chapter 1 that the frontal eye field (FEF) sends input to the superior colliculus (SC); and it was discovered that FEF activity is modulated by expected reward (Glaser et al., 2016). On the other hand, the lateral intraparietal cortex (LIP) is also known to encode value-based decisions in action-based decisions, and it is reported to encode salience (Leathers and Olson, 2012). This region is also known to respond to the novelty of a stimulus (Foley et al., 2014). Therefore, it will have to be determined whether movement vigor is decided by value or salience (or both).

We examined these questions by measuring saccades in a task where subjects learned to associate value with neutral abstract stimuli paired with different amounts of gains or losses. In decision trials, which involved risky options, they chose an option with a saccade to maximize reward, from which we inferred the utility of individual options. In probe trials, subjects made a cued saccade in the direction of a single stimulus, from which we measured vigor and reaction time. We found that subjects showed faster movement and quicker reaction toward a stimulus that was associated with gain, and the opposite for a stimulus that was associated with loss. More notably, we observed across-subject differences in evaluation and behavior and it showed a strong association with subjective value.

4.2. Materials and Methods

4.2.1. Participants and general apparatus

Subjects (Experiment 1: n=24, 26.3±8.2 years old, mean±SD, 8 females; Experiment 2: n=22, 26.3±7.5 years old, mean±SD, 10 females) sat in a well-lit room in front of an LED monitor (59.7 x 33.6 cm, 2560 x 1440 pixels, light gray background, frame rate 144 Hz) placed at a distance of 35 cm. The subject's head was restrained using a bite bar. They viewed images (4x4°, except as noted) and we measured their eye movements using an EyeLink 1000 (SR Research) infrared recording system (sampling rate 1 kHz). Only the right eye was tracked. All subjects were naïve to the paradigm. The experiments were approved by the Johns Hopkins University School of Medicine Institutional Review Board, and all subjects signed written consent form approved by the board. Subjects were paid $15/hour regardless of any behavioral outcome.

4.2.2. Behavioral task – Experiment 1

Our goal was to assess whether the saccade vigor of an individual was associated with the value that they had learned to assign to an abstract stimulus. To assess subjective value, we performed an experiment in which people learned the values of 10 abstract visual stimuli through experience. Each stimulus was a 2° x 2° colored square, designated with a "+" or "-" (Fig. 4-1B). Each square was randomly assigned to a point distribution, with a mean that ranged from a loss of 5 points to a gain of 5 points. The points associated with each color were selected randomly in each trial from a beta distribution with parameters $\alpha = \beta = 2$, scaled so that each color box was associated with a single mean: -5, -4, …, +5. The plus and minus indicator at the center of the

square noted the sign of the mean of the distribution. The color to point association was selected randomly for each subject but remained consistent throughout the experiment. For example, the plus yellow square in Fig. 4-1B was associated with a distribution with mean equal to a gain of 4 points, and the minus yellow square was associated with mean equal to a loss of 4 points. In addition to these 10 colored squares, a black square with "0" at the center was associated with exactly 0 points.

The experiment contained two types of trials, randomly intermixed. Both types of trials (Fig. 4-1A) began with a center fixation period that lasted for 1 sec and ended with a beep (1 kHz). In decision trials, the fixation point was replaced with three different colored stimuli. One stimulus appeared alone and represented a sure bet (100% probability of acquiring the points associated with that stimulus). The other two stimuli appeared together and represented a risky bet (each with 50% probability). The participant had 5 seconds to indicate their choice by making a saccade to one of the two stimuli (square, 0.5° x 0.5°) that appeared on the horizontal axis at ±20°. Once the saccade concluded, the stimuli at the center were erased and the trial consequences were displayed for 1 sec: the earned stimulus was displayed at the dot location along with text that indicated the number of points acquired. The points were drawn from the random distribution associated with the colored stimulus. Failure to make a choice within the time limit resulted in a loss of 10 points. The trial ended with the display of the color stimulus and the amount of points gained or lost for that trial (duration of 1 sec).

In probe trials, the fixation point was removed, a single stimulus (chosen at random from the 10 colored stimuli) was displayed at the center, and a dot appeared on the horizontal axis (at ±20°). This was the instruction for the subject to make a saccade to the dot. Once the saccade concluded, the stimulus at the center was erased and displayed at the dot location, along with text

that indicated the number of points that the subject had gained or lost for the trial. As in the decision trials, the points were drawn from the random distribution associated with the colored stimulus.

Before the start of the experiment, the subjects were instructed that there were 10 stimuli consisting of two sets of five colored boxes that represented points that could be gained or lost in each trial. “Each color will indicate how many points you will gain or lose. The black box will always give zero points when chosen. Boxes with plus signs will add to your score, while boxes with minus signs will decrease your score. For example, if an orange box with a plus sign indicates a gain of 10, an orange box with a minus sign will indicate a loss of 10.”

The experiment consisted of 11 blocks, each with 100 trials. The first block was a training block and began with 100 points and included only probe trials. The remaining 10 blocks each had 40 probe trials and 60 decision trials, distributed randomly. The total score was reset to 100 at the start of the second block. For probe trials, each of the 10 colored squares was presented with equal frequency within each block, distributed randomly. In probe trials, the direction of the dot was chosen randomly on the left or right but with equal frequency within each block.

For a decision trial, we randomly picked three stimuli from among the 11 stimuli. We presented the medium-valued stimulus as the sure bet and the other two stimuli as the risky bet. Subjects were not provided any information about the value of the stimuli and thus had to make their decisions solely based on the consequences of previous trials. The side that represented the sure bet was random and chosen with equal left-right frequency for each block. Following completion of the second block, the final score of the previous block was carried over as the

starting score of the next block. At the conclusion of every fourth trial, the total score earned was displayed at center fixation.

4.2.3. Behavioral task – Experiment 2

This version of the experiment used the same set of visual stimuli and structure of blocks and trials. It only differed from the previous experiment in how the probe trials were presented.

In probe trials, instead of presenting a single stimulus, two stimuli were presented, one on each side (Fig. 4-4A). One of the stimuli was always the black box stimulus. Instead of giving the "go" signal simultaneously with the display of stimuli, we instructed the subjects to wait for ~1 second (0.8~1.2 second, uniform random distribution) after the two stimuli were displayed. After this time, we eliminated one of the stimuli and displayed a dot on the horizontal axis (at ±20°). Subjects were instructed to move their eyes toward the dot. The rest of the procedure was the same as in Experiment 1. The chance of which stimuli should be selected, and chance of each side was equal for both options available. The chance of each stimulus (that is not the black box) appearing was also equal for all options.

Decision trials were given in the same manner, again randomly intermixed with probe trials, again in the same manner.

4.2.4. Data analysis

Eye position data were filtered with a third-order Savitzky-Golay filter (frame size 11). Saccade onset and offset were determined in real time with 20°/s threshold. We identified valid saccades as those that occurred between stimuli with start and endpoints that were within 5° of the boundaries of the start and end images (to account for the fact that the subjects were not

specifically instructed to fixate on a precise location). For probe trials, we excluded reaction times that were longer than 1 sec.

Our objective was to test whether behavior in probe trials reflected the subjective value that we had estimated from decision trials. Thus, we analyzed the vigor of saccades only in probe trials, and inferred subjective value based on choices made in decision trials. Statistical testing relied on linear mixed-effect models. In each model, the dependent variables were saccade peak velocity and reaction time, fixed effects were stimulus objective value or subjective value, and random effects were individuals. Dependent variables were normalized for each individual by dividing the measured value by the within-subject mean. Statistics were performed on normalized dependent variables.

Once we determined the reaction time and peak saccade velocity associated with a given stimulus in the probe trials, we asked whether vigor could serve as a proxy for subjective value. To evaluate the accuracy of such a policy, for each stimulus in the probe trials we computed saccade velocity and reaction time, imagined that the subjective value of the stimulus was set by these variables, and then used these measures in each decision trial to predict choice. For example, to evaluate the vigor policy, we assigned subjective value to the 11 stimuli based on vigor in the probe trials, and then used this to predict the choice of the participant in each of the decision trials: pick the option that has the larger vigor estimated subject value (100% of the vigor for the sure option stimulus, vs. the sum of 50% of vigor for each of the risky option stimuli). We compared the accuracy of this vigor-based and reaction time-based policy with a policy that made choices based on subjective values that were estimated based on the actual decisions of each subject. We also estimated the accuracy of a reward-based policy, calculated in the same manner, but excluding trials with equal expected reward. To predict outcomes, we used

the actual options faced by each participant. We used a Wilcoxon signed-rank test to compare the performance of the various policies.

4.2.5. Estimating subjective value of stimuli

The objective value of each stimulus was set by the mean of the point distribution associated with each colored square (Fig. 4-1B). The participants formed subjective values based on the choices that they made and the points that they earned. We estimated the subjective value of a stimulus for each subject based on their choices in the decision trials.

In a decision trial, the choice was between a sure option (a single stimulus) and a risky option (two stimuli, 50% chance of each). To model the choices that subjects made, we designed a one-layer perceptron network that had as its input the three stimuli that were available in each trial. The output of the network was the probability of picking the sure option (Fig. 4-1C). Input $\boldsymbol{x}$ was an 11-element vector, with each element representing one of the stimuli $x_1, \ldots, x_{11}$ starting from the most negative to the most positive, and with the black square (0 points) being the sixth element. In each trial, the input vector $\boldsymbol{x}$ was set so that one element had a value of -1 for the sure stimulus, two elements had a value of 0.5 for the pair of risky stimuli, and 0 for the remaining elements. The weight vector $\boldsymbol{u}$ represented the subjective value of each stimulus and was also an 11-element vector. A linear combination of the available stimuli was represented with variable z:

$$z = \boldsymbol{u}^{\top}\boldsymbol{x} \tag{4.1}$$

For example, if in a given trial the sure option was stimulus x_4, and the risky option was stimuli x_2 and x_7, then $z = 0.5(x_2 + x_7) - x_4$. In other words, the variable z represented the difference between the subjective values of the two options. This was then transformed via a

logistic function that produced an output y that represented the probability of picking the sure option:

$$y = \frac{1}{1 + \exp(-z)} \tag{4.2}$$

From the decisions that each participant made we estimated the subjective value that they had assigned to each stimulus, represented by the weight vector $\boldsymbol{u}$. We assumed that the subjective value of the zero stimulus (the sixth element of $\boldsymbol{u}$) was exactly zero. To find the remaining weights, we used a binary cross-entropy loss function:

$$\text{loss} = -\frac{1}{N}\sum (t \log y + (1 - t)\log(1 - y)) \tag{4.3}$$

In the above equation, N is the total number of decision trials (600). Binary variable t represented the actual decision of the subject, with 1 for choosing the risky option, and 0 for the sure option. To find $\boldsymbol{u}$, we differentiated Eq. (3) with respect to $\boldsymbol{u}$, thus providing a stochastic gradient descent estimate of the subjective values. We stopped the algorithm when the norm of change of the subjective value $\Delta\boldsymbol{u}$ was less than 10^{-4}, within 10,000 iterations.

4.3. Results

4.3.1. Subjective value reflects choice behavior

In this task we had two types of trials. In decision trials, people choose one of two options, a sure option and a risky option. In probe trials, people are given the cue to choose the only option present. These two types of trials are related to the subjective values of the stimuli. Thus, it is important to estimate the utilities of each color box accurately so that we can argue

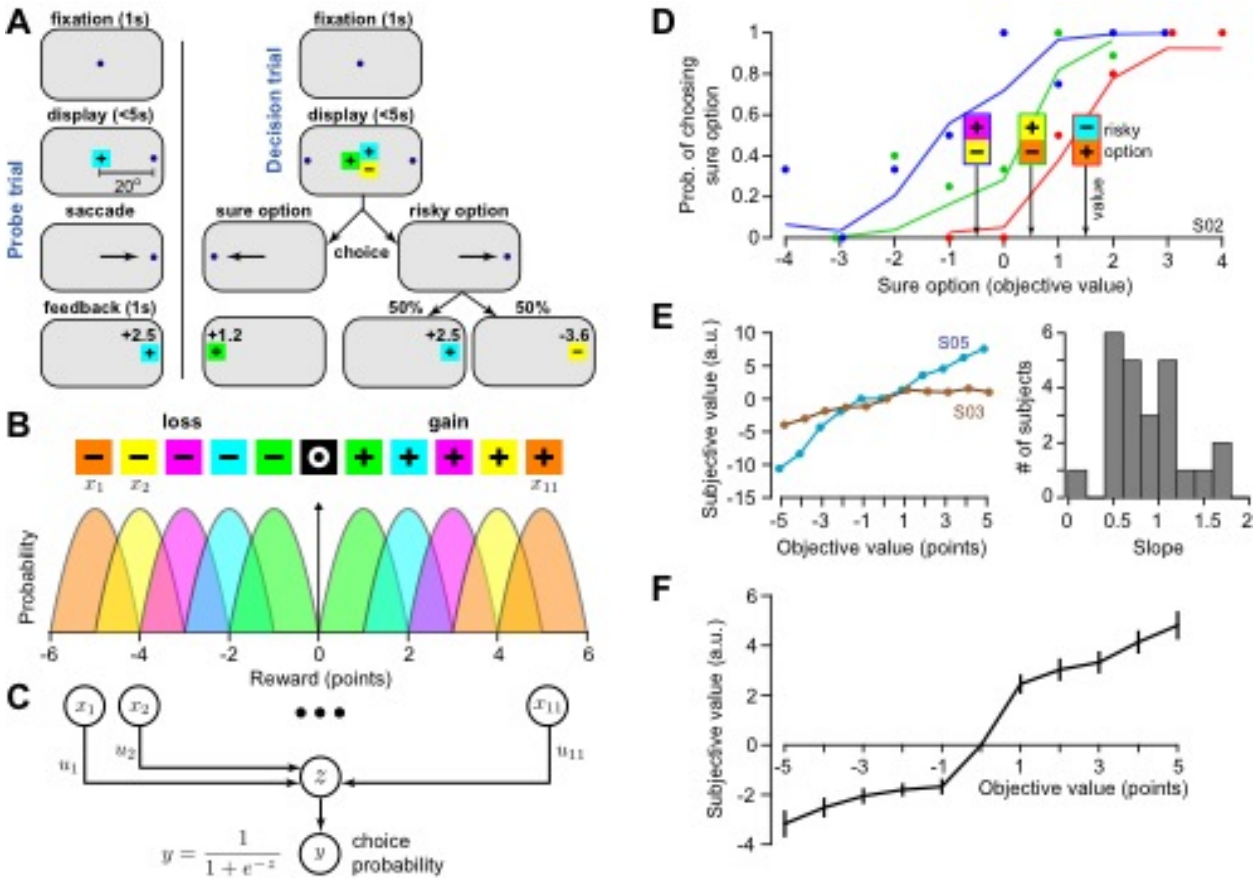

Figure 4-1. Utilities are derived from choice behavior. (**A**) Experimental protocol for experiment 1. In probe trials, subjects are cued to choose the only stimulus displayed. In decision trials, subjects choose either the sure option or the risky option (50% of each color box). (**B**) Example of probability distribution of points from each color box. The distributions are randomly assigned for each subject but are the same throughout the entire session. Colors encode the magnitude of points, and signs inside the boxes encode gain/loss. (**C**) Single-layer perceptron network designed to learn utility. Inputs represent what stimuli have appeared in that trial, and weights are utilities. The intermediate value is the difference between the utilities of the two options (risky minus sure). This value is sent to a logistic activation function to get choice probability. (**D**) Actual and estimated choice probabilities for example pairs for subject S02. The output of the neural network fits the behavior well. (**E**) Example of a utility curve by expected value of point for each color box (left). Each individual shows different utility curves. The slope distribution (right) shows the variety of utility curves. (**F**) Utility curve averaged across all subjects. Error bars represent mean ± SEM across all subjects.

that there is a significant relationship between subjective value and how people move towards visual stimuli associated with reward.

To actually relate movement behavior to utility, inferring subjective value is a very important part of the analysis. Due to the limitations of the present methods to infer utility (explained in the discussion section in detail), an efficient way to figure out utility from a relatively few, simple trials was needed. Therefore, we adopted a neural network to figure out utility (Fig. 4-1C). All subjects but one (who we excluded from all analyses due to very low performance) converged in experiment 1(all subjects converged in experiment 2). For these subjects, a utility curve could be obtained (Fig. 4-1E-F). Since every person was fit with the same logistic function (Eqn. 4.2), utility values refer to how much a subject prefers the option over a default stimulus of zero utility (in this task, the black box).

The subjects show a variety of utility curves, but the utility values are closely related to the objective reward ($p < 10^{-30}$). People show differences in the steepness of the curve (0.88±0.40, mean±SD, Fig. 4-1E). The prediction rate of the utility-based policy (see methods for how policies are applied) on actual choice behavior is better than the prediction rate of the reward-based policy (Fig. 4-3D, $p = 1.12 \times 10^{-4}$ for easy choices, $p = 4.02 \times 10^{-5}$ for hard choices). For this and future comparisons, we classified decision trials into easy trials (reward difference ≥ 1) and hard trials (reward difference < 1). It could also be observed that it fits the behavior of individual trials well (Fig. 4-1D). We claim that we acquired utility values that are the best values to explain people's choice behavior.

4.3.2. Effect of gain and loss on vigor and reaction time

In probe trials, people's motor performance in the task has little to do with their final score; the reward itself is random in a fixed distribution associated with the given stimuli. Nevertheless, saccade vigor and reaction time (RT) are both still modulated by reward. For

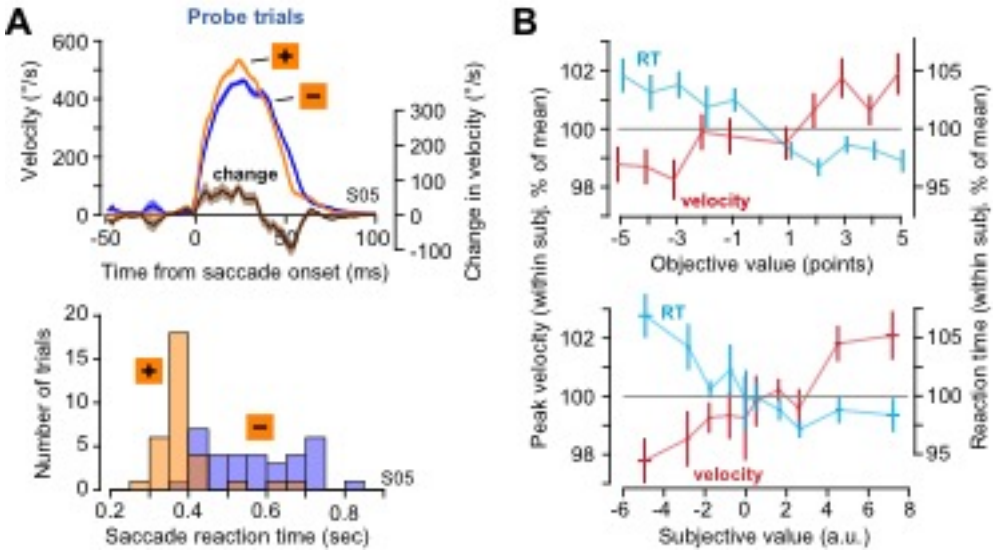

Figure 4-2. Saccade vigor and reaction time is decided by value, not salience. (**A**) Velocity trace of saccades (top) and histogram of reaction time (bottom) for an example pair of stimuli for subject S05. Even with the same number of points, the person shows faster saccade and quicker reaction to gain. (**B**) Saccade vigor and reaction time by reward value (top) and subjective value (bottom). Both show significant effects of values, and all of them are monotonic, and not U-shaped. Error bars for all plots represent mean ± SEM across all subjects.

example, if we compare between the highest-reward stimulus (mean=5) and the lowest-reward stimulus (mean=-5) (for example, the orange color set in Fig. 4-1B), the higher reward stimulus has higher vigor and quicker reaction in probe trials (Fig. 4-2A). This is only representative of the effect of gain and loss, but it implies that reward fastens movement and quickens reaction.

A closer look at the overall data on all stimuli show that this is generally true. Linear model analysis reveals the effect of signed reward value on both behavioral metrics (Fig. 4-2B, top; vigor: $p < 10^{-6}$, RT: $p < 10^{-10}$). This effect states that saccade vigor is not affected by salience, which would be represented by the color and not by the sign on the color boxes, and is

instead significantly affected by the reward value. The effect on RT is the opposite, such that it is quicker for higher-valued stimuli, but it still reaches the same conclusion.

It was also observed that the subjective values acquired from decision trials and the neural network (Fig. 4-1F) have a strong relationship with vigor and RT (Fig. 4-2B, bottom; vigor: $p < 10^{-7}$, RT: $p < 10^{-12}$). However, since utility and objective reward have a very strong relationship, this is expected. One interesting thing to note here, though, is how large the effect size is for vigor and RT. An additional linear mixed-effect model with RT reversed around 1 (so, RT -> 2 minus RT), so that both vigor and RT increase by value, revealed a relative effect size on vigor and RT via its interaction term. In the loss domain, the effect is larger on RT ($p < 10^{-3}$); in the gain domain, the effect is larger on vigor ($p = 0.036$). This reflects the asymmetry of the utility curve, which has a wider range, and implies that vigor might be a better representation of utility, although this is likely task-specific since we do not see the loss aversion commonly reported in decision tasks.

4.3.3. Subjective value determines vigor and reaction time

Subjects show different patterns of utility (Fig. 4-1E-F), and it is an interesting question whether subjects who assign larger values to a given stimulus also show higher vigor than others. To look into this question more closely, we constructed a graphical model that describes how values could influence decision and movement (Fig. 4-3A). The objective value perceived by the sensory system will be converted into a subjective value, and this will decide choice behavior. In our hypothesis (H_1), we suggest that objective value is converted to subjective value, which decides both choice behavior and movement vigor. On the other hand, in an alternative null

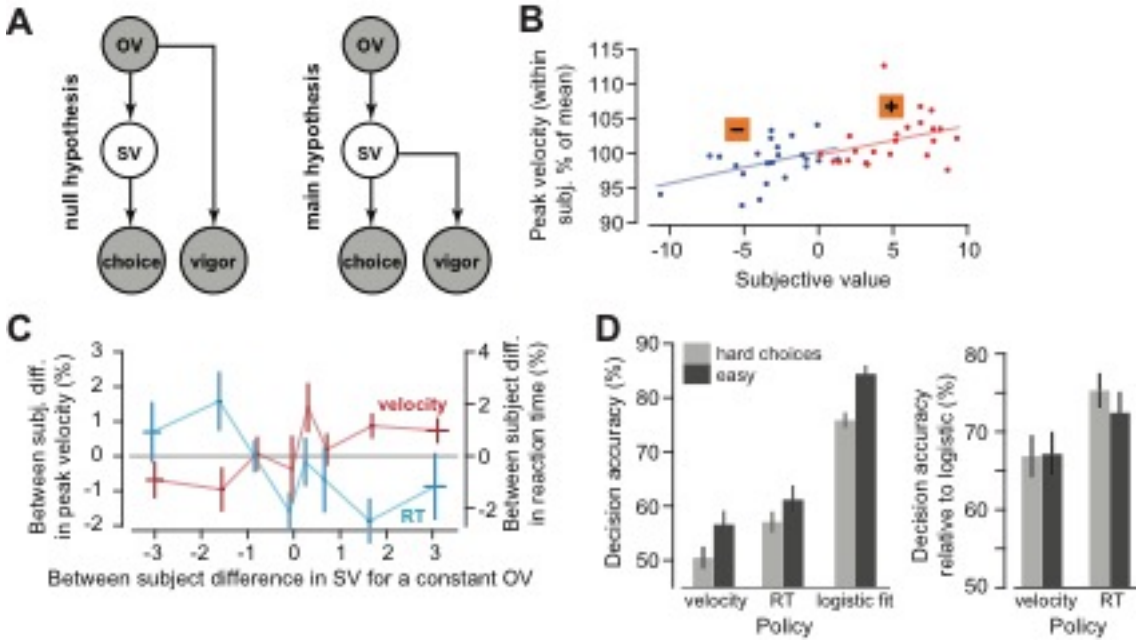

Figure 4-3. Subjective evaluation directly governs movement vigor and reaction. (**A**) Hypothesis of the quantity that guides vigor. The main hypothesis is that subjective value decides both choice and vigor. The null hypothesis directs objective value (points) to modulate vigor, without any direct link between subjective evaluations. (**B**) Example of how individual differences in subjective value result in differences in the peak velocity of saccades, for an example pair. (**C**) Deviation of saccade vigor (normalized for each subject) by deviation of subjective value. Saccade vigor and reaction time still shows the same tendency as they did in Figure 4-2, even with constant objective reward. (**D**) Prediction accuracy of policies depending purely on control of movement (vigor, RT). Although fitting by utility (logistic fit) trivially gives the best prediction rate, velocity and RT policies both give better results than choosing randomly (50%). We also see that predictions are better for easy trials than for hard trials. Error bars for all plots represent mean ± SEM across all subjects.

hypothesis (H_0), vigor would only be decided by the objective reward of the target, independent of any influence of their subjective evaluation, while decisions are still done by utilities.

To test this hypothesis, we looked at the individual stimulus and at the data among subjects (Fig. 4-3B). In this example, some subjects have higher utility values, while some have lower utility values, even within the distribution for the same objective reward. Moreover, there

looks to be a trend between utility assignment and vigor: subjects that assign a higher utility value to the same stimulus exerts higher vigor toward that stimulus.

To test whether this effect is statistically prevalent, we applied the linear mixed-effect model to the "deviation" of vigor and "deviation" of utility. In this analysis, deviation refers to individual value minus the mean value across subjects for a particular stimulus the individual value is assigned to. This test effectively assessed how between-subject evaluation is related to the between-subject difference in utility. The model gives a significant result for both vigor (Fig. 4-3C, $p = 0.0047$) and reaction time (Fig. 4-3C, $p = 0.0036$). Thus, we keep our main hypothesis (H_1) and conclude that vigor and reaction time are affected by subjective evaluation even when the objective reward stays the same.

Finally, we analyzed how well vigor and reaction time fit people's choice behavior. To assess this, we suggested alternative "policies" that subjects could implement to make decisions. Section 4.3.1 already describes how good the utility-based policy is in predicting choice behavior for both hard trials and easy trials. For a policy to be relevant in predicting choice, it should be better than the prediction rate of a random policy, which is simply choosing any option with 50% probability. A random policy will always have a 50% prediction rate for choice behavior.

For hard choices, the vigor-based policy showed no difference from 50% (Fig. 4-3D, left, $p = 0.99$). However, the RT-based policy showed a significantly better prediction rate than the random policy (Fig. 4-3D, left, $p = 0.0042$). The relative accuracy of the RT-based policy compared to the utility-based policy was about 75.3±2.2% (mean±SEM) (Fig. 4-3D, right). For easy choices, both vigor-based and RT-based policies were significantly better in predicting choices than random policy (Fig. 4-3D, left; vigor: $p = 0.0225$, RT: $p < 10^{-3}$). The relative

accuracies of these cases were about 67.1±2.9% and 72.4±2.7% (both mean±SEM) for each (Fig. 4-3D, right). Thus, we observed the relevance of these policies in predicting decisions.

4.3.4. Modulation of vigor and response time during decision making

So far, we have analyzed people's behaviors in probe trials and their relationship with utility from decision trials. We also analyzed movement velocity and decision time in decision trials, since we think they would be also modulated by values of the stimuli. We hypothesized that either the difference between the expected utilities of two options or the sum of expected utilities of two options (or both) would affect vigor and decision time.

We saw effects on vigor and reaction time regardless of the parameter selected between reward and utility. Movement vigor is modulated not by the difference in expected reward/utility (Fig. 4-4A), but the sum (Fig. 4-4B, reward: $p < 10^{-3}$, utility: $p < 10^{-5}$), such that the eye is faster when the overall value of the trial is higher. This actually corresponds to our earlier finding in Chapter 2 (Fig. 2-2B). Meanwhile, decision time is affected by both the difference in expected reward/utility (Fig. 4-4A, reward: $p < 10^{-7}$, utility: $p < 10^{-17}$) and the sum (Fig. 4-4B, reward: $p < 10^{-30}$, utility: $p < 10^{-24}$). In detail, decision time is shorter when the difference is larger (it is easier to choose the better option) and when the overall value of the trial is higher. Not only probe trials but also the decision trials show modulation of movement.

4.3.5. Reward prediction error and its potential effect

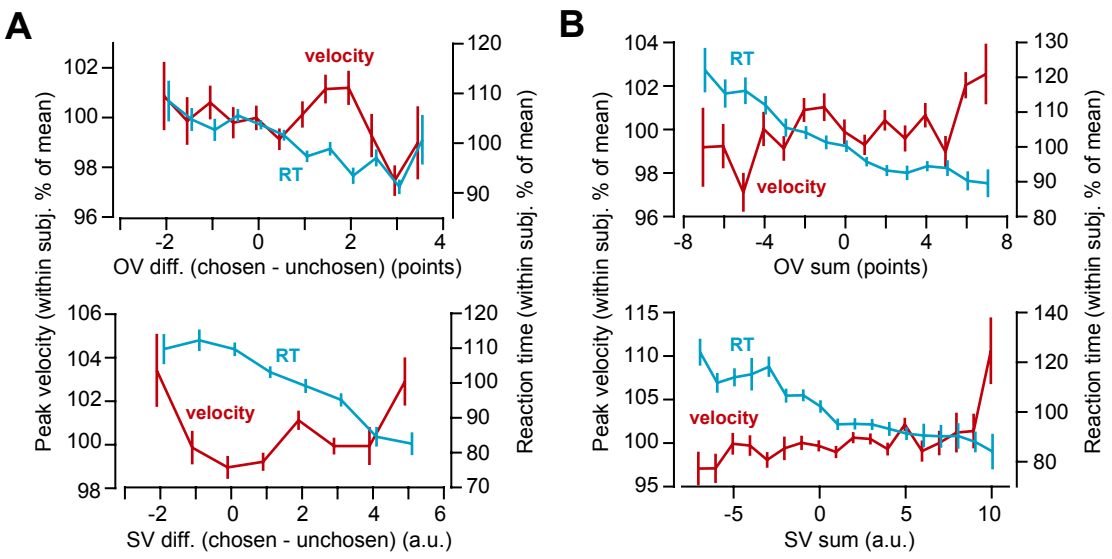

Figure 4-4. Control of movement during decision. (**A**) Modulation of saccade vigor and reaction time by difference between objective values (top) and subjective values (bottom). Velocity is not affected by relative difference value. (**B**) The same plots but by the sum of objective/subjective values. The sum of values affects both variables. Error bars for all plots represent mean ± SEM across all subjects.

Since dopamine neurons encode reward prediction error, we proposed the possibility that it is the reward prediction error, instead of reward, that determines movement vigor. To test this hypothesis, we conducted Experiment 2, which had a behavioral task very similar to experiment 1, but with one major change (Fig. 4-5A). In this experiment, we changed the probe trials so that it briefly showed (~1s) two visual stimuli, one box on each side, before removing one of them, and the subjects were to make a saccade toward the target dot 20° away in the direction of the remaining stimulus. Basically, the role of the 1s window display period was to have subjects form a belief about the expectation of an outcome; we set an equal chance of a color box on each side remaining, thus having the expected value be the mean of the values of both stimuli. One of the targets was always the black box, and thus 50% of the probe trials were those made toward

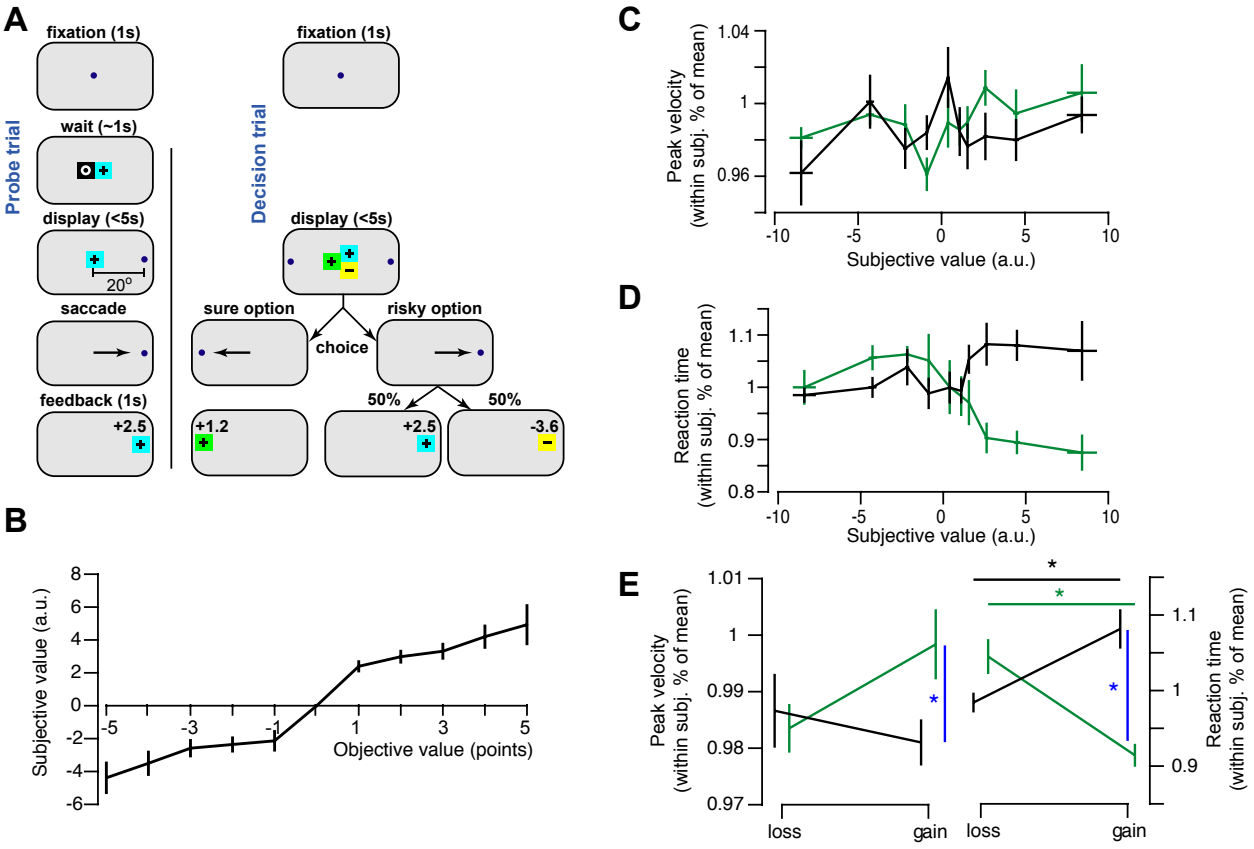

Figure 4-5. Reward prediction errors affect movement but in a limited way. (**A**) Experimental protocol for experiment 2. In this experiment, probe trials now display two color boxes, but the subjects always have to wait until the go cue, which points to one of the stimuli. (**B**) Utility curve across all subjects. This curve is derived from decision trials in the same way as in experiment 1 (see methods and Figure 4-1). (**C**) Vigor in probe trials by utility of stimulus. It is separated by whether the cue was toward the non-black box (green) or toward the black box (black). (**D**) The same plot as in (C), but for reaction time. (**E**) Comparison between gain and loss stimuli on saccade vigor and RT. Different lines are for the same standards as in Figures (C) and (D). (*p<0.05) Error bars for all plots represent mean ± SEM across all subjects.

the non-black stimulus, and the other 50% were those made toward the black box. Utility analysis was made in the same way in Experiment 1 (Fig. 4-5B).

For each type of probe trial, we analyzed the effect of the subjective value of the non-black stimulus on vigor and RT (Fig. 4-5C-D). Most of the linear mixed-effect models do not give any significant effects of utility on saccade vigor or reaction time, except the effect on

reaction time for probe trials with saccades made toward the non-black stimulus (Fig. 4-5D, green line, $p = 0.0027$). Binary comparison between stimuli with negative reward and positive reward separately gives more insight. For effects on vigor, there is a significant difference between probe trial types of gain stimuli (Fig. 4-5E, left, $p = 0.016$). For effects on RT, not only is there a significant difference between probe trial types of gain (Fig. 4-5E, right, $p < 10^{-3}$), but there is also a significant difference between gain and loss for both probe trial types (Fig. 4-5E, right, toward non-black stimulus: $p < 10^{-3}$, toward black box: $p < 10^{-3}$). The rest of the binary comparisons had no effect.

4.4. Discussion

The brain makes decisions based on the subjective values of available options. However, the subjective value is a hidden value that we cannot directly measure. We've already observed instances of vigor modulation in the previous chapters. In this chapter, we look at what exact quantity (such as utility, salience, points) movement vigor could serve as a proxy.

Here, we use abstract visual stimuli associated with the fixed distribution of gain or loss. Subjects performed two types of trials intermixed: decision trials, from which we inferred the utility subjects assigned to each stimulus, and probe trials, from which we observed motor behavior. We found that vigor and reaction time varied monotonically with the value of the stimulus; vigor was high and RT was short for a stimulus that forecasted high gain, and the opposite for a stimulus that forecasted high loss. Vigor and RT showed differing degrees of sensitivity in each gain or loss domain. From the probe trials, we think that vigor is modulated by reward value rather than salience.

Furthermore, a critical question is whether the between-subject difference in evaluation causes a difference in vigor. We found that for a given objective reward, there was a relationship between utility and vigor/RT. Individuals who assigned a greater value to a stimulus also moved faster and reacted quicker to that stimulus. Also, we observed that the prediction rate of behavior-based policy is significantly better than chance in most cases, suggesting that choice behavior could be predicted from motor behavior to a certain degree.

To precisely assess the effect of subjective value, inferring utility values from decision trials is crucial. Unfortunately, utility is a subjective value that is hard to measure. Studies in neuroscience that aim to estimate utility use a concept called certainty equivalent (CE), which is a quantity that has the same utility with the option of choice. If the option is a risky one that has a 50% probability of resulting in one of two values, the utility of a value of the CE will be the mean value of the utilities of the two values. This mathematical property makes estimating utility possible by measuring the CE for various options. CE could simply be directly judged by the subjects (Grether and Plott, 1979), but better alternative methods are fitting a logistic function between the fixed risky option and the variable sure option, or even using a psychological adaptive method such as PEST (parameter estimation by sequential testing), which proved to be better (Bostic et al., 1990; Christopoulos et al., 2009; Stauffer et al., 2014).

Even under the oversimplified assumption that utility is invariant throughout the experiment (at least after the training block) and that subjects make rational decisions to maximize reward, utility is still hard to measure accurately for two reasons: it is subjective, which means that it should be measured independently for each subject, thus drastically increasing the number of choices needed for each subject, and it is not directly estimated from choice, so that a choice does not give an expected utility value for each option, thus requiring a

large number of choices to make an indirect estimate. These problems have been partly solved by the methods mentioned above, but it still requires many repetitions, while adding another limitation that if a similar decision is presented repetitively many times, individuals will make decisions based on their memory instead of evaluating utility independently each time for each trial. Thus, we wanted an efficient way to infer utility from a relatively smaller number of trials, for a limited number of associated stimuli, with random presentation, without needing to change subsequent trials based on subjects' choice.

We implemented a neural network, which is very efficient to "learn" utilities by an optimization algorithm such as gradient descent. We believe that our estimates of utility values are the quantities that best explain the choices of each subject across all decision trials, since they are the values with minimum loss. The high accuracy of prediction (Fig. 4-3D) to the real decision indicates that this value is accurate to predict choice behavior. We only used 600 trials for each subject, which is a small number, with all stimuli chosen at random, and its randomness allowed subjects to judge utility in every trial instead of referring to their previous choice by memory.

The effect of utility could be predicted by neurophysiological evidence. The superior colliculus is known to send input to bursting neurons that send initiation signals to motor neurons. Not only this, but it also controls saccade vigor and reaction time (Dorris et al., 1997; Smalianchuk et al., 2018). The superior colliculus receives excitatory inputs from the cortical areas, such as the frontal eye field (FEF) and the lateral intraparietal area (LIP), both of which house neurons that respond stronger to stimuli that predict greater reward (Glaser et al., 2016; Platt and Glimcher, 1999). The superior colliculus also receives inhibitory input from the basal ganglia, specifically the substantia nigra pars reticulata (SNr), which houses neurons that also

respond to associated rewards (Sato and Hikosaka, 2002). Thus, it is likely that the subjective value of a visual stimulus could be reflected in the change of input signals from these areas, which in turn would influence behavior so that greater value results in faster movement and quicker reaction.

However, collicular activity and cortical and basal ganglial inputs are also affected by stimulus salience, which makes it difficult to predict what determines vigor and RT. Neurons in the superior colliculus show modulation by salience based on trial context (Kim and Basso, 2010). The LIP neurons also show activity modulation by salience or novelty rather than reward value (Leathers and Olson, 2012; Foley et al., 2014). In the basal ganglia, the SNr is modulated by the striatum in various pathways, and dopamine neurons regulate how these striatal neurons respond to cortical inputs. It was also shown that while some dopamine neurons show greater response to higher reward, some others show greater response to both higher reward and punishment (Matsumoto and Hikosaka, 2009).

Our results show undisputedly that saccade vigor is a reflection of subjective value and not salience. However, it is still possible that if negative feedback exists in a different form (such as an air puff), it could result in increased cortical and dopaminergic activity which would result in greater vigor.

Our results also show that movement vigor is an indicator of actual subjective value, rather than simply a reflection of sensory input of objective value. Our hypothesis (H_1) is validated if the across-subject difference in subjective value reflects the across-subject difference in vigor. Not only this, but previous studies have shown that lack of vigor sensitivity to reward coincides with lack of sensitivity of dopamine neurons and caudate nuclei activity (Kawagoe et

al., 2004). Thus, these findings suggest that behavioral results could provide a real-time proxy for subjective evaluation.

Decision trials are harder to analyze due to the limited number of trials compared to the variety of combinations of the stimuli displayed, compared to probe trials (3 vs. 1). However, we saw modulation of both vigor and decision time. Vigor increases with the value of both options, which is interesting since it is faster even when the non-chosen option has higher value. In fact, this observation corresponds to our results in Chapter 2 (Fig. 2-2B), where we saw that vigor between two images was decided by both the target and the starting image in a positive relationship. Maybe vigor is a reflection of the overall value of all available options. The fact that vigor does not rely on the difference between the two options implies that it is not always the same as the target value.

We also observed that decision time is modulated by both the difference and the sum of the values of the two options. Common models such as the drift diffusion model (DDM) predict this effect in a two alternative forced choice (2AFC) task (Bogacz et al., 2006). Also, it has been observed that reaction time is quicker when the sum of the values of two options is higher (Bari et al., 2019). Our analysis of decision time also revealed similar effects. It is interesting that as in this example, vigor and decision time do not always follow each other, implying a difference in nature between the two variables.

Unfortunately, in experiment 2, we could not find any evidence that reward prediction error could also modify movement vigor. However, the fact that we could not observe the effects seen in experiment 1 suggests that the effects are more reflective in nature. The prediction error we are controlling in experiment 2 is not the error between the prediction and actual reward of

the same associated visual stimulus, however, so it also suggests that the effects could be tied to each individual stimulus already associated with a certain utility value.

5. Conclusion and General Discussion

This thesis has presented how the control of movement is closely related to decision making. People cannot consciously decide their own saccade velocity, but interestingly, many studies have shown that it is still modulated by reward. The close relationship between the saccadic circuitry and decision-making pathway suggests that people control saccade vigor in a different way based on reward and effort, on which they make decisions. Our primary question is how saccade velocity is modulated by factors that affect decision making in a quantifiable way.

In Chapter 2, we showed how saccade vigor is modulated in a two-target gaze task. When participants could decide how long they gazed at each stimulus, they preferred certain types of images and positions with small eccentricity, suggesting that eccentricity modulates the effort required to hold the eyes on a fixed location. We also observed that saccade vigor increases as both images become more preferred, regardless of whether it was at the start or the destination. This fact led us to consider the richness of the environment as the factor that decides saccade vigor.

In Chapter 3, we suggested a refined version of the marginal value theorem (MVT) to explain both harvest time and movement vigor. We created a task environment that closely resembled an actual foraging situation yet was significantly simplified, and thus were able to observe the effects of past, present, and future reward and effort on gaze time as they would be from the predictions of the model. We also observed movement vigor being modulated by past and future effort. We suggested an alternative explanation of how saccade vigor is modulated by effort justification.

In Chapter 4, we tested a hypothesis about how vigor could reflect the subjective value of a stimulus. We designed a task so that it contained both decision trials, from which we assessed choice behavior and inferred utility, and probe trials, in which we observed saccade vigor and reaction time. The results show that participants display faster saccades and quicker reaction towards targets with higher utility, regardless of their salience, which is the magnitude of gain or loss in this task. We also made a between-subject analysis and showed that the effects are affected by subjective value even with the stimulus of the same reward.

All these chapters describe the modulation of peak velocity of saccades in decision making. As mentioned, the movement "module" of eyes has advantages in simplification, but due to its nature, several aspects cannot be tested. For example, we described saccade velocity being modulated by the effort required to hold the eyes, but it is unclear whether this applies to other forms of effort, such as movement effort, due to our inability to directly modulate the difficulty of eye movement (for example, we cannot apply a force field). In general, it would be interesting to test whether the same phenomena are observed for other movement modules, such as arm movements, walking, et cetera. We mention a few studies that showed that other types of movements are also modulated by reward (Rigoux and Guigon, 2012; Summerside et al., 2018). Other movement modules generally have more complicated neural circuits for the control of movement and involve many brain areas and muscles, but they enable scientists to test more various theories.

While all chapters point to the fact that movement vigor is modulated by reward and effort, this is confined to simple task structures. Our current version of the theory is a simplified version and is difficult to apply to a more general situation. Optimal control theory in motor system is a general theory that explains how actions are selected and performed to minimize a

cost function given some constraints. In an invariant situation, an organism is expected to make a movement that achieves their goal with lowest cost, usually in the form of integrated jerk (Flash & Hogan, 1985). However, this accounts for neither uncontrolled trial-to-trial variability (Scholz & Sch ner, 1999) nor compensation for perturbations (Robertson & Miall, 1997). Thus, the theory has recently been referred to as optimal feedback control (OFC) (Todorov & Jordan, 2002). According to recent studies, an organism constructs an internal model of the body and the environment; this is formed by past motor experience and the resulting feedback, including continuous online feedback to modulate the motion already being performed (K rding and Wolpert, 2006; Diedrichsen et al., 2010; Shadmehr et al., 2010).

Although the concept of optimal feedback control and modulation of movement by reward seems closely related, the research that integrates between the two is only at its initial step. It has been observed that feedback based on reward prediction errors can also induce motor learning (Izawa & Shadmehr, 2010; Wolpert & Landy, 2012). Some studies suggest models that integrate a temporally discounted reward with motor cost to explain how organisms change movement vigor in the presence of reward (Rigoux & Guigon, 2012; Shadmehr et al., 2016). Currently, most theories are on the stage of development, and is yet to be fully integrated with optimal feedback control, if possible.

In conclusion, these findings suggest that control of movement is significantly affected by components that decide humans' choice behaviors. This work also has implications for the neural circuitry that governs movement vigor and decision making in humans. The various areas that are related to one of the two aspects affect each other, and our behavioral observations confirm the firing patterns and connections of these areas. Future work on movement vigor and

its relationship to decision making will further enlighten our knowledge of the neural basis of both and hopefully facilitate related studies between the two fields of study.

Bibliography

Allais, M., Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'ecole americaine, *Econometrica*, 21(4), 503-546 (1953)

Amiez, C., Joseph, J-P., and Procyk, E., Anterior cingulate error-related activity is modulated by predicted reward, *Eur. J. Neurosci.,* 21(12), 3447-3452 (2005)

Aronson, E., and Mills, J., The effect of severity of initiation on liking for a group, *J. Abnorm. Soc. Psychol.,* 59(2), 177-181 (1959)

Bahill, A.T., Clark, M.R., and Stark, L., The main sequence, a tool for studying human eye movements, *Mathematical Biosciences*, 24, 191-204 (1975)

Bannerman, R.L., Milders, M., de Gelder, B., and Sahraie, A., Orienting to threat: faster localization of fearful facial expressions and body postures revealed by saccadic eye movements, *Proc. Royal Soc. B: Biol. Sci.,* 276, 1635-1641 (2009)

Barack, D.L., Chang, S.W.C., and Platt, M.L., Posterior cingulate neurons dynamically signal decisions to disengage during foraging, *Neuron*, 96(2), 339-347 (2017)

Bari, B., Grossman, C.D., Lubin, E.E., Rajagopalan, A.E., Cressy, J.I., and Cohen, J.Y., Stable representations of decision variables for flexible behavior, *Neuron*, https://doi.org/10.1016/j.neuron.2019.06.001 (2019)

Bayer, H.M., and Glimcher, P.W., Midbrain dopamine neurons encode a quantitative reward prediction error signal, *Neuron*, 47(1), 129-141 (2005)

Beckstead, R. M., Domesick, V. B., and Nauta, W. J. H., Efferent connections of the substantia nigra and ventral tegmental area in the rat, *Brain Res.,* 175(2), 191-217 (1979)

Beeler, J.A., Daw, N.D., Frazier, C.R.M., and Zhuang, X., Tonic dopamine modulates exploitation of reward learning, *Front. Behav. Neurosci.,* 4:170 (2010)

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., and Cohen, J., The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks, *Psych. Rev.,* 113(4), 700-765 (2006)

Bostic, R., Herrnstein, R.J., and Luce, R.D., The effect on the preference-reversal phenomenon of using choice indifferences, *J. Econ. Behav. Org.,* 13, 193-212 (1990)

Brodeur, M.B., Dionne-Dostie, E., Montreuil, T., and Lepage, M., The Bank of Standardized Stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research, *PLoS ONE*, 5(5), e10773 (2010)

Cannon, S.C., and Robinson, D.A., Loss of the neural integrator of the oculomotor system from brain stem lesions in monkey, *J. Neurophys*., 57(5), 1383-1409 (1987)

Charnov, E.L., Optimal foraging: the marginal value theorem, *Theor. Popul. Biol*., 9, 129-136 (1976)

Choi, J.E.S., Vaswani, P.A., and Shadmehr, R., Vigor of movements and the cost of time in decision making, *J. Neurosci.,* 34(4), 1212-1223 (2014)

Christopoulos, G.I., Tobler, P.N., Bossaerts, P., Dolan, R.J., and Schultz, W., Neural correlates of value, risk, and risk aversion contributing to decision making under risk, *J. Neurosci.,* 29(40), 12574-12583 (2009)

Clement, T.S., Feltus, J.R., Kaiser, D.H., and Zentall, T.R., "Work ethic" in pigeons: Reward value is directly related to the effort or time required to obtain the reward, *Psychon. Bull. Rev.,* 7(1), 100-106 (2000)

Cohen, J.Y., Amoroso, M.W., and Uchida, N., Serotonergic neurons signal reward and punishment on multiple timescales, *eLife*, 4:e06346 (2015)

Cools, A. R., Role of the neostriatal dopaminergic activity in sequencing and selecting behavioural strategies: Facilitation of processes involved in selecting the best strategy in a stressful situation, *Behav. Brain Res.,* 1(5), 361-378 (1980)

da Silva, J.A., Tecuapetla, F., Paixão, V., and Costa, R.M., Dopamine neuron activity before action initiation gates and invigorates future movements, *Nature*, 554, 244-248 (2018)

Diedrichsen, J., Shadmehr, R., and Ivry, R.B., The coordination of movement: optimal feedback control and beyond, *Trends in Cog. Sci.,* 14(1), 31-39 (2010)

Dorris, M.C., Paré, M., and Munoz, D.P., Neuronal activity in monkey superior colliculus related to the initiation of saccadic eye movements, *J. Neurosci.,* 17(21), 8566-8579 (1997)

Ellsberg, D., Risk, ambiguity, and the savage axioms, *Quart. J. Econ.,* 75(4), 643-669 (1961)

Fischer, B., and Boch, R., Saccadic eye movements after extremely short reaction times in the monkey, *Brain Research*, 260(1), 21-26 (1983)

Flash, T., and Hogan, N., The coordination of arm movements: an experimentally confirmed mathematical model, *J. Neurosci.,* 5(7), 1688-1703 (1985)

Foley, N.C., Jangraw, D.C., Peck, C., and Gottlieb, J., Novelty enhances visual salience independently of reward in the parietal lobe, *J. Neurosci.,* 34(23), 7947-7957 (2014)

Gerfen, C. R., Engber, T. M., Mahan, L. C., Susel, Z., Chase, T. N., Monsma Jr., F. J., and Sibley, D. R., D1 and D2 dopamine receptor-regulated gene expression of striatonigral and striatopallidal neurons, *Science*, 250, 1429-1432 (1990)

Gerfen, C. R., Herkenham, M., and Thibault, J., The neostriatal mosaic: II. Patch- and matrix-directed mesostriatal dopaminergic and non-dopaminergic systems, *J. Neurosci.,* 7(12), 3915-3934 (1987)

Glaser, J.I., Wood, D.K., Lawlor, P.N., Ramkumar, P., K rding, K.P., and Segraves, M.A., Role of expected reward in frontal eye field during natural scene search, *J. Neurophys*., 116(2), 645-657 (2016)

Grether, D.M., and Plott, C.R., Economic theory of choice and the preference reversal phenomenon, *Amer. Econ. Rev.,* 69(4), 623-638 (1979)

Haith, A., Reppert, T. R., and Shadmehr, R., Evidence for hyperbolic temporal discounting of reward in control of movements, *J. Neurosci.,* 32 (34), 11727-11736 (2012)

Handel, A., and Glimcher, P.W., Quantitative analysis of substantia nigra pars reticulata activity during a visually guided saccade task, *J. Neurophys*., 82(6), 3458-3475 (1999)

Hayashi, K., Nakao, K., and Nakamura, K., Appetitive and aversive information coding in the primate dorsal raphé nucleus, *J. Neurosci.,* 35(15), 6195-6208 (2015)

Hayden, B.Y., Pearson, J.M., and Platt, M.L., Neuronal basis of sequential foraging decisions in a patchy environment, *Nat. Neurosci.,* 14, 933-939 (2011)

Heimer, L., The olfactory cortex and the ventral striatum, *Limbic Mechanisms,* 95-187 (1978)

Hikida, T., Kimura, K., Wada, N., Funabiki, K., and Nakanishi, S., Distinct roles of synaptic transmission in direct and indirect striatal pathways to reward and aversive behavior, *Neuron*, 66(6), 896-907 (2010)

Hikosaka, O., and Wurtz, R. H., Visual and oculomotor functions of monkey substantia nigra pars reticulata. IV. Relation of substantia nigra to superior colliculus, *J. Neurophys.,* 49(5), 1285-1301 (1983)

Ilango, A., Kesner, A. J., Keller, K. L., Stuber, G. D., Bonci, A., and Ikemoto, S., Similar roles of substantia nigra and ventral tegmental dopamine neurons in reward and aversion, *J. Neurosci.,* 34(3), 1703-1713 (2014)

Itoh, H., Nakahara, H., Hikosaka, O., Kawagoe, R., Takikawa, Y., and Aihara, K., Correlation of primate caudate neural activity and saccade parameters in reward-oriented behavior, *J. Neurophys*., 89(4), 1774-1783 (2003)

Izawa, J., and Shadmehr, R., Learning from sensory and reward prediction errors during motor adaptation, *PLoS Comp. Bio.,* 7, e1002012 (2011)

Jóhannesson, Ó., and Kristjánsson, Á., Violating the main sequence: asymmetries in saccadic peak velocities for saccades into the temporal versus nasal hemifields, *Exp. Brain. Res.,* 227(1), 101-110 (2013)

Kacelnik, A., and Marsh, B., Cost can increase preference in starlings, *Anim. Behav.,* 63(2), 245-250 (2002)

Kahneman, D., and Tversky, A., Prospect Theory: An analysis of decision under risk, *Econometrica*, 47(2), 263-291 (1979)

Kawagoe, R., Takikawa, Y., and Hikosaka, O., Expectation of reward modulates cognitive signals in the basal ganglia, *Nat. Neurosci*., 1, 411-416 (1998)

Kawagoe, R., Takikawa, Y., and Hikosaka, O., Reward-predicting activity of dopamine and caudate neurons—a possible mechanism of motivational control of saccadic eye movement, *J. Neurophys.,* 91(2), 1013-1024 (2004)

Keller, E.L., Participation of medial pontine reticular formation in eye movement generation in monkey*, J. Neurophys*., 37(2), 316-332 (1974)

Kim, B., and Basso, M.A., A probabilistic strategy for understanding action selection, *J. Neurosci.,* 30(6), 2340-2355 (2010)

Kobayashi, S., Nomoto, K., Watanabe, M., Hikosaka, O., Schultz, W., and Sakagami, M., Influences of rewarding and aversive outcomes on activity in macaque lateral prefrontal cortex, *Neuron*, 51(6), 861-870 (2006)

K rding, K.P., and Wolpert, D.M., Bayesian decision theory in sensorimotor control, *Trends in Cog. Sci.,* 10(7), 319-326 (2006)

Kravitz, A. V., Tye, L. D., and Kreitzer, A. C., Distinct roles for direct and indirect pathway striatal neurons in reinforcement, *Nat. Neurosci.,* 15, 816-818 (2012)

Leathers, M.L., and Olson, C.R., In monkeys making value-based decisions, LIP neurons encode cue salience and not action value, *Science*, 338, 132-135 (2012)

Liu, Z., Luo, P., Wang, X., and Tang X., Deep learning face attributes in the wild, *The IEEE Int. Conf. on Comp. Vis. (ICCV),* 3730-3738 (2015)

MacArthur, R.H., and Pianka, E.R., On optimal use of a patchy environment, *The American Naturalist*, 100, 603-609 (1966)

Manohar, S.G., Chong, T.T.-J., Apps, M.A.J., Batla, A., Stamelou, M., Jarman, P.R., Bhatia, K.P., and Husain, M., Reward pays the cost of noise reduction in motor and cognitive control, *Curr. Biol*., 25(13), 1707-1716 (2015)

Matsumoto, M., and Hikosaka, O., Two types of dopamine neuron distinctly convey positive and negative motivational signals, *Nature*, 459, 837-841 (2009)

Matsumoto, M., Suzuki, W., and Tanaka, K., Neuronal correlates of goal-based motor selection in the prefrontal cortex, *Science*, 301, 229-232 (2003)

Mimura, M., Oeda, R., and Kawamura, M., Impaired decision-making in Parkinson's disease, *Parkinsonism & Related Disorders*, 12(3), 169-175 (2006)

Niv, Y., Daw, N.D., Joel, D., and Dayan, P., Tonic dopamine: opportunity costs and the control of response vigor, *Psychopharmacology*, 191(3), 507-520 (2007)

O'Doherty, J., Winston, J., Critchley, H., Perrett, D., Burt, D.M., and Dolan, R.J., Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness, *Neuropsychologia*, 41(2), 147-155 (2003)

Öhman, A., Flykt, A., and Esteves, F., Emotion drives attention: Detecting the snake in the grass, *J. Exp. Psychol.: General*, 130(3), 466-478 (2001)

Parthasarathy, H.B., Schall, J.D., and Graybiel, A.M., Distributed but convergent ordering of corticostriatal projections: analysis of the frontal eye field and the supplementary eye field in the macaque monkey, *J. Neurosci*., 12(11), 4468-4488 (1992)

Pasquereau, B., and Turner, R.S., Limited encoding of effort by dopamine neurons in a cost–benefit trade-off task, *J. Neurosci*., 33(19), 8288-8300 (2013)

Platt, M.L., and Glimcher, P.W., Neural correlates of decision variables in parietal cortex, *Nature*, 400, 233-238 (1999)

Pollack, I., Methodological examination of the PEST (Parametric Estimation by Sequential Testing) procedure, *Perception & Psychophysics*, 3(4), 285-289 (1968)

Reppert, T.R., Lempert, K.M., Glimcher, P.W., and Shadmehr, R., Modulation of saccade vigor during value-based decision making, *J. Neurosci*., 35(46), 15369-15378 (2015)

Reppert, T.R., Rigas, I., Herzfeld, D.J., Sedaghat-Nejad, E., Komogortsev, O., and Shadmehr, R., Movement vigor as a traitlike attribute of individuality, *J. Neurophys.,* 120(2), 741-757 (2018)

Rigoux, L., and Guigon, E., A model of reward- and effort-based optimal decision making and motor control, *PLoS Comp. Biol.,* 8(10), e1002716 (2012)

Robertson, E. M., and Miall, R. C., Multi-joint limbs permit a flexible response to unpredictable events, *Exp. Brain Res.,* 117(1), 148-152 (1997)

Robinson, D.A., Oculomotor unit behavior in the monkey, *J. Neurophys.,* 33(3), 393-403 (1970)

Robinson, D.A., Models of the saccadic eye movement control system, *Biol. Cybernetics*, 14(2), 71-83 (1973)

Robinson, D.A., and Fuchs, A.F., Eye movements evoked by stimulation of frontal eye fields, *J. Neurophys*., 32(5), 637-648 (1969)

Roesch, M. R., Taylor, A. R., and Schoenbaum, G., Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation, *Neuron*, 51(4), 509-520 (2006)

Sato, M., and Hikosaka, O., Role of primate substantia nigra pars reticulata in reward-oriented saccadic eye movement, *J. Neurosci.,* 22(6), 2363-2373 (2002)

Schelp, S.A., Pultorak, K.J., Rakowski, D.R., Gomez, D.M., Krzystyniak, G., Das, R., and Oleson, E.B., A transient dopamine signal encodes subjective value and causally influences demand in an economic context, *Proc. Natl. Acad. Sci. of USA (PNAS),* 114(52), E11303-11312 (2017)

Schlag, J., and Schlag-Rey, M., Evidence for a supplementary eye field, *J. Neurophys*., 57(1), 179-200 (1987)

Scholz, J. P., and Sch ner, G., The uncontrolled manifold concept: identifying control variables for a functional task, *Exp. Brain Res.,* 126(3), 289-306 (1999)

Schultz, W., Dayan, P., and Montague, P.R., A neural substrate of prediction and reward, *Science*, 275, 1593-1599 (1997)

Sedaghat-Nejad, E., Herzfeld, D.J., and Shadmehr, R., Reward prediction error modulates saccade vigor, *J. Neurosci.,* 39(25), 5010-5017 (2019)

Segraves, M.A., Activity of monkey frontal eye field neurons projecting to oculomotor regions of the pons, *J. Neurophys*., 68(6), 1967-1985 (1992)

Shadmehr, R., Huang, H.J., and Ahmed, A.A., A representation of effort in decision making and motor control, *Curr. Biol.,* 26, 1929-1934 (2016)

Shadmehr, R., Orban de Xivry, J. J., Xu-Wilson, M., and Shih, T.-Y., Temporal discounting of reward and the cost of time in motor control, *J. Neurosci.,* 30(31), 10507-10516 (2010)

Shadmehr, R., Reppert, T.R., Summerside, E.M., Yoon, T., and Ahmed, A.A., Movement vigor as a reflection of subjective economic utility, *Trends in Neurosci.,* 42(5), 323-336 (2019)

Shadmehr, R., Smith, M. A., and Krakauer, J. W., Error correction, sensory prediction, and adaptation in motor control, *Ann. Rev. Neurosci.,* 33, 89-108 (2010)

Smalianchuk, I., Jagadisan, U.K., and Gandhi, N.J., Instantaneous midbrain control of saccade velocity, *J. Neurosci.,* 38(47), 10156-10167 (2018)

Stauffer, W.R., Lak, A., and Schultz, W., Dopamine reward prediction error responses reflect marginal utility, *Curr. Biol.,* 24(21), 2491-2500 (2014)

Strassman, A., Highstein, S. M., and McCrea, R. A., Anatomy and physiology of saccadic burst neurons in the alert squirrel monkey. I. Excitatory burst neurons, *J. Comp. Neurol.,* 249(3), 337-357 (1986a)

Strassman, A., Highstein, S. M., and McCrea, R. A., Anatomy and physiology of saccadic burst neurons in the alert squirrel monkey. II. Inhibitory burst neurons, *J. Comp. Neurol.,* 249(3), 358-380 (1986b)

Summerside, E.M., Shadmehr, R., and Ahmed, A.A., Vigor of reaching movements: reward discounts the cost of effort, *J. Neurophys.,* 119(6), 2347-2357 (2018)

Sutton, R.S., Learning to predict by the methods of temporal differences, *Machine Learning*, 3(1), 9-44 (1988)

Sylvestre, P.A. and Cullen, K.E., Quantitative analysis of abducens neuron discharge dynamics during saccadic and slow eye movements, *J. Neurophys.,* 82(5), 2612-2632 (1999)

Tachibana, Y., and Hikosaka, O., The primate ventral pallidum encodes expected reward value and regulates motor action, *Neuron*, 76(4), 826-837 (2012)

Takikawa, Y., Kawagoe, R., Itoh, H., Nakahara, H., and Hikosaka, O., Modulation of saccadic eye movements by predicted reward outcome, *Exp. Brain Res.,* 142(2), 284-291 (2002)

Todorov, E., and Jordan, M. I., Optimal feedback control as a theory of motor coordination, *Nat. Neurosci.,* 5, 1226-1235 (2002)

Ungelstedt, U., Stereotaxic mapping of the monoamine pathways in the rat brain, *Acta Physiologica Scandinavica*, 82(S367), 1-48 (1971)

Van Gisbergen, J.A., Robinson, D.A., and Gielen, S., A quantitative analysis of generation of saccadic eye movements by burst neurons, *J. Neurophys*., 45(3), 417-442 (1981)

Vergilino-Perez, D., Fayel, A., Lemoine, C., Senot, P., Vergne, J., and Doré-Mazars, K., Are there any left-right asymmetries in saccade parameters? Examination of latency, gain, and peak velocity, *Invest. Ophthal. Vis. Sci.,* 53(7), 3340-3348 (2012)

von Neumann, J.V., and Morgenstern, O., *Theory of Games and Economic Behavior*, Princeton, NJ: Princeton University Press (1944)

Watabe-Uchida, M., Zhu, L., Ogawa, S. K., Vamanrao, A., and Uchida, N., Whole-brain mapping of direct inputs to midbrain dopamine neurons, *Neuron*, 74(5), 858-873 (2012)

Wolpert, D. M., and Landy, M. S., Motor control is decision-making, *Curr. Opin. Neurobiol.,* 22(6), 996-1003 (2012)

Xu-Wilson, M., Zee, D.S., and Shadmehr, R., The intrinsic value of visual information affects saccade velocities, *Exp. Brain Res*., 196(4), 475-481 (2009)

Yoon, T., Geary, R.B., Ahmed, A.A., and Shadmehr, R., Control of movement vigor and decision making during foraging, *Proc. Natl. Acad. Sci. of USA (PNAS),* 115(44), E10476-10485 (2018)

Young, W. S. III, Alheid, G. F., and Heimer, L., The ventral pallidal projection to the mediodorsal thalamus: a study with fluorescent retrograde tracers and immunohistofluorescence, *J. Neurosci.,* 4(6), 1626-1638 (1984)

Zarrugh, M.Y., Todd, F.N., and Ralston, H.J., Optimization of energy expenditure during level walking, *Eur. J. Appl. Physiol. Occup. Physiol.,* 33, 293-306 (1974)

Zentall, T.R., Animals prefer reinforcement that follows greater effort: Justification of effort or within-trial contrast?, *Comp. Cog. Behav. Rev.,* 8, 60-77 (2013)

Curriculum Vitae

Tehrim Yoon, B.S.

Updated September, 2019

Educational History:

Ph.D. expected	2019	Department of Neuroscience Johns Hopkins School of Medicine Mentor: Reza Shadmehr, Ph.D. Laboratory for Computational Motor Control
B.S.	2014	Department of Physics and Astronomy (physics major) Department of Chemistry – double major Seoul National University

Other Professional Experience

Research associate	2013-2014	Laboratory for physics of biological systems, Seoul National University Mentor: Moo-Young Choi, Ph.D.
Research associate	2012	Laboratory for Bioinformatics, Seoul National University Mentor: Sung-Roh Yoon, Ph.D.

Scholarships, fellowships, or other external funding

2014-2019 Samsung Scholarship: Doctoral Study Overseas Program

2011-2014 Korea Presidential Science Scholarship Award

Academic and other honors

2014 *Summa cum laude* Seoul National University

2010 Gold Medal International Physics Olympiad (IPhO)

Publications, peer reviewed, including reviews

Shadmehr R, Reppert TR, Summerside EM, **Yoon T**, Ahmed AA (2019) Movement Vigor as a Reflection of Subjective Economic Utility. *Trends in Neurosciences* 42(5):323-336

Yoon T, Geary RB, Ahmed AA, Shadmehr R (2018) Control of movement vigor and decision making during foraging. *Proceedings of National Academy of Sciences (USA)* 115(44):E10476-E10485

Publications, chapters and other non-peer reviewed

Invited talks

Movement vigor and its implications in subjective value for saccades. *31st Association for Psychological Sciences (APS) Annual Convention*, Washington D.C., May 2019.

Poster, abstracts, etc.

Yoon T, Shadmehr R (2019) Saccade vigor as an implicit measure of subjective value. *Soc. Neurosci.,* Chicago, IL, October 2019 (accepted)

Yoon T, Shadmehr R, Ahmed AA (2018) Movement vigor and decision-making in a patchy reward environment. *Soc. Neurosci.,* San Diego, CA, November 2018

Yoon T, Shadmehr R (2017) Interaction between effort and reward in the oculomotor system. *Soc. Neural Control of Movement,* Dublin, Ireland, May 2017

Yoon T, Shadmehr R (2016) Interaction between effort and reward in the oculomotor system. *Soc. Neurosci.,* San Diego, CA, November 2016

Inventions, Patents, Copyrights

Service and leadership, including teaching or tutoring activities

Spring 2017 Teaching assistant for Learning Theory (580.491, 580.691), Johns Hopkins University